A Review of Navigation Algorithms for Unmanned Aerial Vehicles Based on Computer Vision Systems

Article *in* Gyroscopy and Navigation · January 2023

DOI: 10.17285/0869-7035.00105

2 authors:

Bushra Ali
National University of Science and Technology MISIS

Rinat Sadekov
National University of Science and Technology MISIS

ISSN 2075-1087, Gyroscopy and Navigation, 2022, Vol. 13, No. 4, pp. 241–252. © Pleiades Publishing, Ltd., 2022.
Russian Text © The Author(s), 2022, published in Giroskopiya i Navigatsiya, 2022, No. 4, pp. 87–105.

A Review of Navigation Algorithms for Unmanned Aerial Vehicles Based on Computer Vision Systems

B. Ali[a], *, R. N. Sadekov[a], **, and V. V. Tsodokova[b], ***

[a] *The National University of Science and Technology MISIS, Moscow, Russia*
[b] *Concern CSRI Elektropribor, JSC, St. Petersburg, Russia*
**e-mail: bushrayehyaali@gmail.com*
***e-mail: r.sadekov@misis.ru*
****e-mail: tsodokova.vv@gmail.com*

Received May 9, 2022; revised December 13, 2022; accepted December 19, 2022

Abstract—The article analyzes the works devoted to outdoor navigation of unmanned aerial vehicles (UAVs) in GNSS-denied environments using computer vision systems (optical range cameras). The algorithms addressed are based on matching of UAV-generated images with the available georeferenced terrain images. Images are matched either pixel by pixel, by their key points, or using neural networks. The stages of each algorithm, as well as the accuracy achieved by the authors and the field data used for testing are considered. The paper concludes with a discussion of the capabilities and limitations of the proposed approaches.

Keywords: unmanned aerial vehicle, autonomous navigation, key point matching, descriptor, detector, deep learning, computer vision, satellite imagery, correlation-extreme systems

DOI: 10.1134/S2075108722040022

INTRODUCTION

Over the past several years UAVs have been used successfully in a variety of applications, among which are irrigation of agricultural fields [1], firefighting [2–4], package delivery [5], search and rescue operations [6, 7], area monitoring [8], and military operations [9]. UAVs have a number of advantages over unmanned ground vehicles: large area coverage, the ability to land at any point of the terrain, high flight speed and, as a consequence, the speed of packages delivery regardless of the road circumstances, etc. These properties make UAVs the best means for accomplishment the above-mentioned tasks.

The UAV position in space is determined by 6 degrees of freedom: coordinates (x, y, z) and orientation angles (ϕ, θ, ψ). The angles of orientation are the angles relative to horizon ϕ, θ and azimuthal angle ψ. The object position is usually determined in the WGS-84 global coordinate system, but local coordinate systems can also be used depending on the source data.

Traditionally, navigation parameters of UAVs in open space are determined by GNSS signals, while orientation parameters are obtained using micromechanical inertial navigation systems (INS). As practice shows [12–14], the GNSS signal is subject to interference, both of natural (trees, mountains, and urban buildings) and artificial (suppression) origin [15–17]. When GNSS signals are lost, estimation of UAV location becomes unreliable. In the absence of direct satellite visibility, the accuracy of navigation parameters can be provided by INS [10, 11], but only on a small-time interval. In addition, UAVs have to solve various “attendant” problems resulting from such scenarios of their application, e.g. return to the base [18] in the event of control signal loss or generation of smoother flight trajectories involving mapping, which is not feasible only on the basis of GNSS information.

In order to partially or completely replace navigation systems that rely on information from GNSS, navigation and control systems using maps as reference information, as well as computer vision and intelligent methods for data processing are being actively developed [19, 20].

Map-aided navigation systems were developed in the second half of the 20th century [21–23]. In Russian literature, map-aided navigation is known as navigation using geophysical fields or correlation-extreme navigation [25, 26]. This name is due to the criterion in the form of a correlation function, which was used when comparing the reference and fragments of measurement data. Correlation-extreme systems were used not only in the cases when a map was available, but when it was unavailable as well [27, 28].

Systems that use information obtained by means of vision implement the approach called vision-based navigation. Such systems are suited for the following applications [19]: visual localization and mapping, obstacle avoidance, and path planning.

In relation to UAVs, the systems of visual localization and map building are subdivided, in turn, into the following [19]: map-based, map-building and mapless systems. The first group makes use of available reference maps; the second one builds maps during the flight, and the third, mapless systems, do not use any maps at all.

The latter include systems that implement algorithms of mapless visual odometry [29]. Systems that generate maps during the UAV flight use map-building simultaneous localization and mapping (SLAM) algorithms [30]. These algorithms do not involve any reference information: position and orientation of the object are determined by a sequence of images. The key idea of SLAM algorithms is to generate a map during the navigation problem solution and to estimate errors in determining navigation parameters by the constructed map. Thus, angular drift can be compensated for by the following techniques: optimization of a trajectory based on its local fragment (bundle adjustment), trajectory closing after its repeated crossing (loop closing), etc. Visual odometry algorithms as well as SLAM algorithms are not considered in this paper since they are mainly used for indoor navigation and, in addition to vision systems, they use information from a variety of other sensors (lidars, depth cameras, etc.).

The aim of this article is to review the available publications on outdoor vision-based UAV navigation. The algorithms for this kind of navigation systems using reference information are analyzed. They are implemented through different approaches to image matching. In our review the works are grouped as follows: Group 1 deals with correlation-extreme approach, Group 2 is concerned with key point matching, and Group 3 makes use of neural networks. The algorithms given are considered in accordance with this classification.

The first group of algorithms determines the mutual orientation and position of images (the reference one and the one generated on board the UAV) by matching whole images using the correlation-extreme approach. For more reliable comparison of images, contours, areas, points, and other elements characteristic of the image, extracted on the basis of both classical and neural network algorithms of image processing can be used.

In the second group of algorithms, the mutual orientation and position of the image generated by the UAV and the reference image are determined by matching key points that are previously extracted with the help of so-called detectors. The key point is defined by its position coordinates in the image and a descriptor, which is a vector that characterizes the uniqueness of the area around the key point. The detectors are constructed using both classical and neural network-based algorithms.

The third group of image matching algorithms uses deep neural networks that are trained with the help of reference images of the terrain. Neural networks trained on reference images are applied both in global navigation—image-based geolocalization and local navigation—image regression. In the first case, UAV location is determined across large areas, covering the territory of a county, state, and country. In the second case, which is obviously closer to the subject of this article, the UAV position and orientation are determined as it flies across a small area, e.g. a residential courtyard or a block of several houses. The methods of the third group are new in the navigation of moving objects.

It should be noted that the existing navigation algorithms based on reference maps are rather science-intensive and complicated, and they may use techniques from different groups, which makes it sometimes impossible to refer the algorithms considered in a work to a particular group. Apart from the vision system, the authors often use information obtained from inertial systems, so that they are also mentioned in some works.

Thus, the works considered in this paper are presented in three sections:

— image-matching navigation algorithms using the correlation-extreme approach;

— image-matching navigation algorithms using key points;

— image-matching navigation algorithms using neural networks.

Each section of the article includes a description of the main idea, analysis of the works, comparative tables with information on the algorithms used and datasets, as well as the accuracy reported by different authors in their publications. In general, each of the solutions described below uses the image provided by the computer vision system (optical range camera) and the terrain map as input data; at the output, a vector of navigation parameters is formed (a specific composition of the state vector is given in the columns *Error, state vector components* of the tables below).

IMAGE-MATCHING NAVIGATION ALGORITHMS USING THE CORRELATION-EXTREME APPROACH

The works considered in this section present algorithms based on the matching of the reference image and the one obtained from the UAV. In this case, no features are extracted from the images; they are matched as a whole, pixel by pixel, based on the correlation-extreme approach, the main idea of which is to find the maximum of the mutual correlation function of the images.

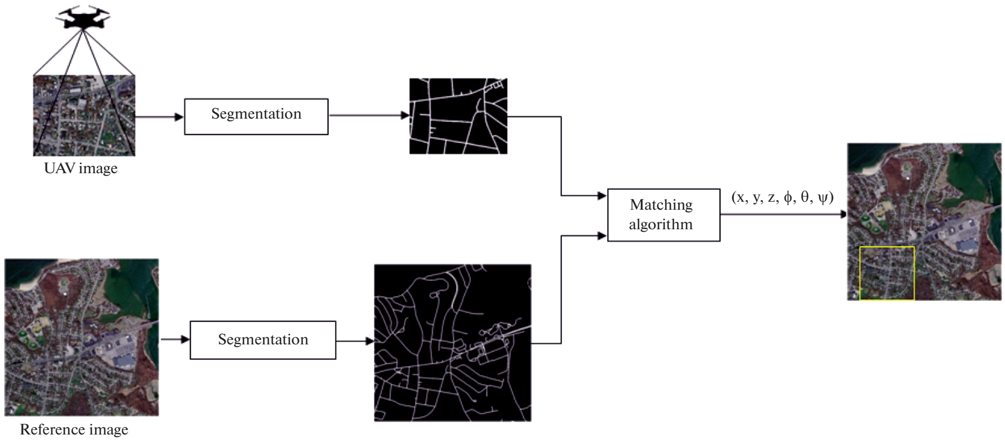

Fig. 1. An image-matching navigation algorithm using the correlation-extreme approach.

The works considered within this group are among the earliest; nowadays the researchers actively use different kinds of neural networks. For example, segmentation neural networks make it possible to use vector maps as reference data, which can be quickly binarized to detect objects on the ground (roads, houses, river networks, etc.).

Figure 1 shows the general stages of the algorithms addressed in this section:

— segmentation aimed to distinguish different classes of objects: roads, buildings, rivers, etc. (for example, in [35] considered below, cGAN neural network was used for segmentation [33, 34]);

— matching of the images obtained;

— additional procedures for correct matching of segmented images, for example, visual odometry algorithms.

In [35], a new approach to visual navigation using conditional generative adversarial neural network cGAN [33, 34] and the correlation-based matching method was presented. The UAV images were segmented and matched with the reference image (reference map) by calculating the correlation based on the normalized sum of squared pixel intensity differences at each time point.

In [36], an approach to UAV navigation based on automatic recognition of landmarks using a correlation approach is described. Images of landmarks with known latitude and longitude information were selected on a reference map. Then, the “candidates” for the desired landmarks were selected on the UAV image using key points and the AKAZE algorithm [37]. The reference landmark image and the images of “candidates” were binarized by means of the Canny edge detector, and the final landmark recognition was performed by binary images, using the mutual correlation function. This approach made it possible to significantly reduce the probability of false landmark identifications.

In [38], the UAV images of buildings of several universities made at different zenith angles of the line of sight were matched with the reference satellite images of these buildings by using CNN neural network. For each building, the University-1652 data set [39] contains one satellite image and 54 UAV images obtained at different zenith angles. CNN was trained in such a way that the UAV image was translated into a form similar to the satellite image, and then they were matched. As a result of the experiments on image matching using CNN, it was possible to achieve a recall level of 90.02% and an average accuracy of 70.82%. It is noteworthy that when matching images, on the contrary, from the satellite view to that from the UAV, the accuracy was reduced by 5%. Though the authors of this work did not estimate navigation accuracy, in their opinion, if necessary, this problem can be solved by adding georeferencing of satellite images.

Conclusions

The algorithms presented in this section are based on image matching (those formed by the UAV and the reference one) using correlation approaches. In order to increase the matching reliability, the image is, first, subjected to segmentation with the use of deep neural networks.

But currently, image matching algorithms based on correlation approaches are used less frequently, giving way to algorithms based on key point matching (see the next section). In addition, despite the active development of neural networks for aerial image segmentation, a large number of object types are still segmented with low accuracy. Note that the works aimed to detect roads and buildings in images, used, first of all, in aerial terrain monitoring or image annotation and in augmented reality systems for UAV pilots are most successful. The high accuracy of segmentation of other types of objects is still limited because of scarcity of labeled data.

Table 1. Comparative characteristics of image-matching navigation algorithms using the correlation-extreme approach

Work	Algorithms	Reference images/Dimension w [m] × l [m]	UAV images/Dimension w [m] × l [m]	Error, state vector components
[35]	cGAN + image correlation	Satellite images/102 × 102	Real flight/216 × 144	AED = 22.7 m
[36]	AKAZE + image correlation	Satellite images (Deimos-2)	Video from the International Space Station (1032 frames)/3840 × 2160	AED = 15.24 m (x, y)
		Satellite images	Real flight	AED = 21.57 m (x, y)
[38]	CNN	University-1652	University-1652	–

Note AED – average Euclidean distance.

IMAGE-MATCHING NAVIGATION ALGORITHMS USING KEY POINTS

The works discussed in this section are devoted to the development of algorithms for matching key points of the reference image and the image from the UAV. Using key points makes it possible to operate with less information for data matching providing high reliability. The basic idea of UAV navigation algorithms using key point matching is shown schematically in Fig. 2; it includes the following stages of the problem solution:

— determination of the UAV initial position (x_0, y_0, z_0, ϕ_0, θ_0, ψ_0), which is implemented if the problem statement requires it. This stage can involve techniques other than key point matching;

— extraction of key points and their descriptors from the image generated by the UAV and the reference image using different point detectors;

— using additional information, e.g., INS measurements or visual odometry algorithms to ensure the continuity of the navigation problem solution in cases where the results of the reference image matching with the UAV image are unsatisfactory;

— preprocessing of the extracted key points aimed to improve them;

— navigation. At this stage, the UAV position in space is determined by matching the key points of the UAV image with those of the reference image, and the transformation matrices of the UAV image coordinate system into the coordinate system of the reference image are determined. This stage can be implemented using either RANSAC methods or a group of the Monte Carlo methods. If fact, both cases are aimed at searching for the maximum mutual correlation function of the key points of the images: the reference one and the one generated by the UAV. Key point matching is carried out by matching their descriptors using the quadratic norm [40] or the Hamming distance (for binary descriptors) [40].

Let us discuss the works presented in the field under consideration from the perspective of the multistage scheme given above.

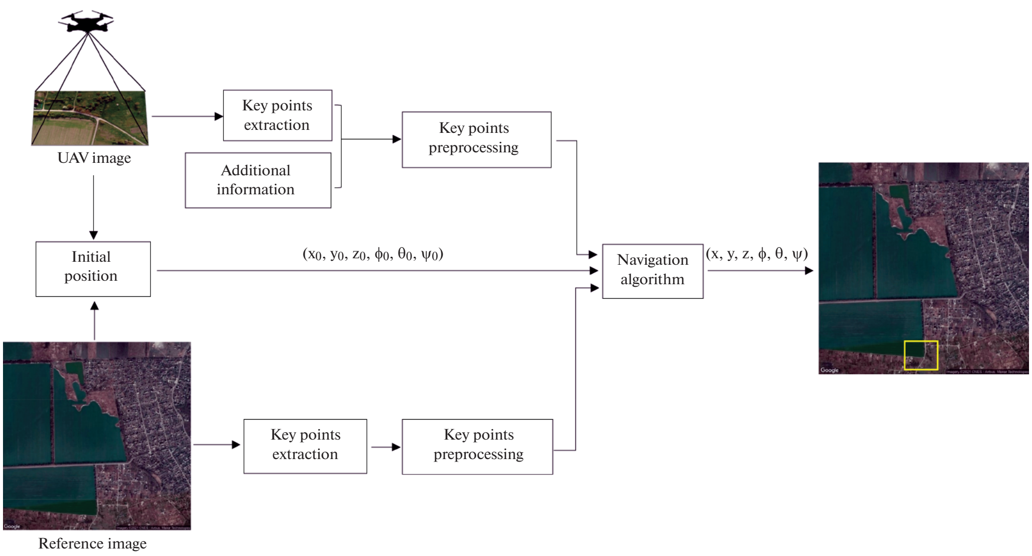

Fig. 2. An image-matching navigation algorithm using key points.

A three-stage algorithm is proposed in [41]. First, the UAV coordinates and orientation parameters estimations were calculated using visual odometry. Then, the UAV images and geo-referenced satellite images were passed through a convolutional neural network composed of VGG16 neural network layers [42] in order to extract the image features for their further matching. A homography matrix was determined using the key points in order to refine the estimated coordinates. The algorithm code is available at https://www.kaggle.com/balraj98/massachusetts-buildings-dataset. The algorithm was tested on two datasets: Village Dataset (https://www.sensefly.com/ education/datasets/?dataset=1419); and Gravel Pit Dataset (https://www.sensefly.com/education/datasets/?dataset=1421).

Similar algorithms for autonomous UAV navigation were used in [46, 47]. SURF algorithm was used as a key point detector [48], while the navigation algorithm was based on RANSAC [43]. Other stages of the scheme described above were not used. Application of RANSAC made it possible to effectively deal with outliers and to find a navigation solution even in noisy conditions. At the same time, the number of key points selected to determine navigation parameters is supposed to be no less than fifteen.

In [49], the extraction of key points was also performed using SURF algorithm. Among those points, ten points with the highest values of the Hessian matrix determinant were selected, for which another descriptor consisting of the responses of various Gabor wavelets was calculated. The algorithm for narrowing the search area was implemented similarly to the one described in [50]. RANSAC was used to match the key points on the reference image and the UAV image.

A modified version of the BRIEF detector [52], abBRIEF, was proposed in [10]. Unlike the original, the descriptor formed by abBRIEF was calculated on the CIELAB color space, and a quantization process was used instead of the Gaussian filter to reduce the image noise. The descriptors formed by abBRIEF are sets (usually 64 or 128) of pairs of pixels that are randomly selected in the vicinity of a key point. Changes of point coordinates between frames were tracked using the optical flow to make matching between the key points during flight simpler. UAV navigation was carried out with the use of the particle filter, in which each particle represented the UAV assumed location. The Hamming distance was used as a measure of the correlation between the key points. Based on the key points found, a homography matrix was calculated, which served for the evaluation of the UAV position in each particle. The datasets POA_UAV and AdM_UAV used in this work are available at https://zenodo.org/record/1244314 and https://zenodo.org/record/1244296, correspondingly.

In [53], an algorithm based on terrain classification is proposed. The following four classes of terrain are distinguished: grass, bushes/trees, street/asphalt, and buildings. The first and second classes refer to areas with complex textures. First, the image was divided into small rectangular areas, and then ORB descriptors were extracted from each area [44]. Then, using the obtained information and the random forest model [54], the areas were classified. The classification allowed for reliable matching of key points by excluding the "problem" areas belonging to the first and second classes. The navigation stage of the algorithm was implemented in the form of the particle filter.

In order to reduce the cost of data storage, a new approach to UAV navigation was proposed in [55]. In the presented algorithm, Google Earth images were used to train an autoencoder, which compressed the image to a lower-dimensional vector representation and decompressed it into the original one during training. The autoencoder is a five-layer CNN encoder and a five-layer CNN decoder. In flight mode, a CNN encoder was used to form a feature vector of UAV images. Then, the extracted vector was matched with the feature vectors of the reference images and the probability of image matching was generated at the output. The location estimate was calculated as the mathematical expectation of the coordinates of the corresponding reference image centers multiplied by the probability of being in those locations.

In [56], an algorithm for UAV navigation using histograms of oriented gradients (HOG) was developed [57]. The first step of the algorithm is to initialize the particle filter in order to obtain the estimated location of the UAV. At this stage, the correlation filter between the UAV image and the reference image, based on 2D Fourier transforms, generated a global plausibility map of the original location, where the location of the maximum was considered as the initial position for the UAV takeoff, with a set of particles formed around it. The next stage used the optical flow [58] and INS measurements to predict the UAV position between frames, which significantly reduced the search space. The HOG descriptors were subsequently extracted from the image generated by the UAV and matched with the point descriptors in the particles localized in the reference image. The experiments showed that the use of the optical flow improved the reliability of the algorithm.

In [59], feature maps for the reference image and the UAV image were generated using ResNet neural network [60, 61]. Then, using the cosine similarity between the extracted features of both images, a 4D correlation map was constructed, from which the indices of matching points were determined and used further in RANSAC (by analogy with the key points) to calculate the transformation matrix. The neural network was adapted and trained using PASCAL-VOC [62] and InLoc [63] datasets. The algorithm code is

Table 2. Comparative characteristics of image-matching navigation algorithms using key points

Work	Algorithms	Reference images/Dimension w [m] × l [m]	UAV images/Dimension w [m] × l [m]	Error, state vector components
[10]	abBRIEF, PF	Google Earth/1160 × 1160	Real flight 3 flights, 2 km	AED = 17.8 m (x, y, z)
[41]	VGG16, VO, SURF	USGS Earth Explorer/ 7582 × 5946	Real flight 0.85 km	AED = 7 m (x, y), 8 m (z) AED = 25 м (x, y), 7.7 m (z)
[46]	SURF, RANSAC	Google Earth/–	Real flight	MAE = 0.11 m, 0.25 m (x, y)
[47]	SURF	Google Earth/3340 × 1880	Simulated from Google Earth/496 × 278	RMSE = 1.6 m (x, y)
[49]	SURF, Gabor wavelets, RANSAC	Satellite image/ 10000 × 10000	Virtual flight	MAE = 4 m (x, y), 0.01° (ϕ, θ), 0.2° (ψ)
[51]	SIFT, ORB, RANSAC, U-Net	Google Maps and Bing	Real flight 1.2 km	AED = 3.6–5.1 m (x, y)
[53]	ORB, random forest, PF	Google Maps/150 × 90	Real flight 60 × 100	AED = 9.5 m (x, y)
[55]	Autoencoder	Google Earth/–	Real flight 1132 m	RMSE = 3 m (x, y), 0.36° (ψ)
[56]	OP, HOG, PF	Google Maps/300 × 150	Real flight 360 m	RMSE = 6.77 m (x, y)
[59]	ResNet, RANSAC	Own dataset	Real flight 3 flights, 0.52 km^2, 0.64 km^2, 0.66 km^2	AED = 3.6 m (x, y)
[64]	U-Net, SIFT, RANSAC	Lyon[5], own dataset'	Real flight 7 km	~100 m (x, y)

Note: OF – optical flow; PF – particle filter; VO – visual odometry; MAE – mean absolute error; RMSE – root-mean-square error; AED – average Euclidean distance.

available at https://github.com/m-hamza-mughal/Aerial-Template-Matching.

In [64], a new approach to the correction of UAV navigation data was proposed based on matching UAV images with the images formed from a vector topographic map. The UAV image was segmented using the U-Net convolutional neural network [31] into two classes: roads/rivers and background. Then, using the SIFT detector [32], key points were extracted from a black-and-white image and matched with the key points of the binary vector-map image. The binary image was formed by setting the white color for roads and hydrographic data (waterbodies). Next, the RANSAC method [43] was used to calculate the homography matrix. The algorithm was tested using vector and raster cartographic data of the city of Lyon (https://rdata-grandlyon.readthedocs.io/en/latest/index.html), as well as the authors' own data, where the results showed the possibility of determining the coordinates with an error of ~100 m.

In [51], to refine the UAV position, in addition to the classical cycle of key point matching, the contours of buildings were extracted from the images (reference one and the one generated by the UAV) using the U-Net neural network [31]. The resulting building contours were evaluated for similarity by morphological features and then used as analogues of key points to calculate the refined homography matrix. The accuracy analysis of the proposed approach was based on two datasets of Famagusta (Cyprus) and Potsdam (Germany).

Conclusions

Table 2 provides details of each paper analyzed above. Some of the approaches use classical key-point extraction algorithms such as SURF [49], ORB [53], and BRIEF [10]. Sometimes, for more accurate matching, images are subjected to a preliminary segmentation [53]. In some works, neural networks are used both for image preprocessing [64] and for extracting their features, which act as key points [59].

Most of the experiments made use of images from real flights to test the approaches, while in [47] the images were borrowed from Google Earth. The results of the works are presented in different metrics (MAE, RMSE, and AED), which, unfortunately, makes it impossible to compare the algorithms from a common

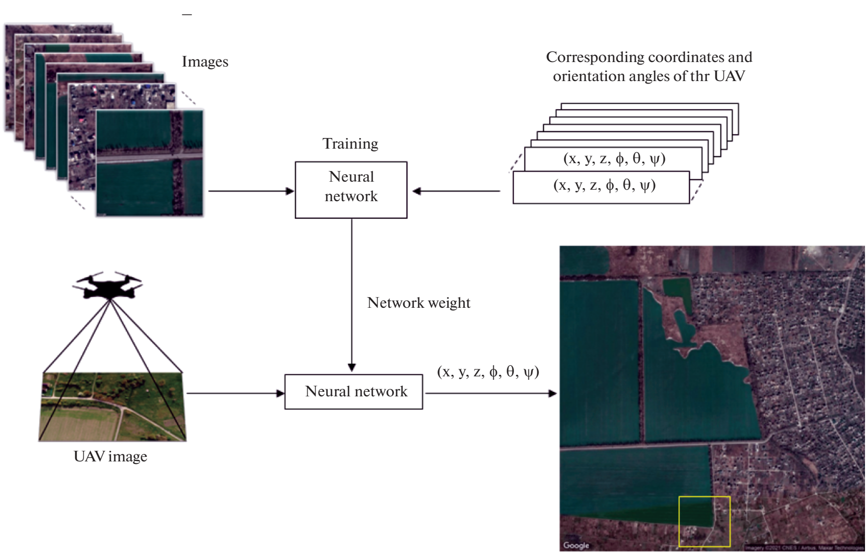

Fig. 3. Image-matching navigation algorithms using neural networks.

position. Also, the authors often give the total (cumulative) accuracy of the obtained solution, rather than separately with respect to each of the state vector components. In addition, some papers use source data that are not available publicly. Therefore, the conclusions are primarily based on the results of "reliable" information.

The large number of works in the field under consideration is indicative of a high demand for the development and practical application of key-point matching navigation algorithms. Accuracy of navigation parameters determined by means of the specified algorithms can vary from tens of meters to decimeters depending on the source data used: images, UAV altitude and speed of flight. In this case, the level of accuracy does not depend on the time and route of the flight. The most accurate results are achieved in the cases where the spatial resolutions of the reference and UAV images match, and also, in when the data is relevant.

In the cases that there is a large amount of geoinformation data, the approaches based on neural networks showed better accuracy and reliability of keypoint extraction compared to classical approaches, which certainly has a positive effect on the quality of the navigation solution. When constructing algorithms, it is often justified to use the optical flow, visual odometry and inertial data, providing high reliability of matching and simplifying the key-point matching procedure.

Such geoinformation resources as Google Earth or Bing Maps can also be used as reference maps in a number of applications, but it is necessary to take into account possible changes in images, which can be caused by various factors: illumination, seasonal and weather phenomena, building development or natural changes in the natural landscape.

IMAGE-MATCHING NAVIGATION ALGORITHMS USING NEURAL NETWORKS

The use of deep learning in computer vision and robotics has increased in recent years with a variety of deep neural network architectures that have proven to be effective in a variety of tasks. One of these cases is the neural network AlexNet [65], which won in a large-scale visual-recognition competition on the ImageNet dataset [66]. Therefore, over the past few years, researchers have been actively using deep learning to implement UAV navigation algorithms.

The algorithms considered in this section are based on a single element – a deep neural network, which, receiving an image as an input, generates the UAV coordinates and orientation parameters at the output. Before the flight, the neural network is trained on a set of available images; as a matter of fact, in this case, the weighted sum of the squared differences in coordinates and orientation parameters is minimized: those "predicted" by the neural network and true ones (corresponding to the set of available images).

The algorithms of this group consist of the following stages, Fig. 3:

— training, at which the neural network is trained on the available georeferenced images or the images for which the object's exact location and orientation are known;

— navigation, at which the UAV location is determined based on network weights obtained at the first stage.

In [67], the architecture of the Xception neural network [68] was modified by adding fully-connected dropout layers and changing the loss function. The negative log likelihood was used as the loss function. Accuracy was estimated using 13 datasets (ADOP 2017 One Foot Ortho (raster) dataset. https://gis.arkansas.gov/; NAIP Imagery dataset. https://www.fsa. usda.gov/programs-and-services/aerial-photography/imagery-programs/naip-imagery/index) from different areas of Washington County (twelve of which were used to train the network). The accuracy analysis showed that the RMSE was 166.5 m in an area of 32.2 km × 32.2 km. The authors studied the accuracy on larger-scale images, but the results proved to be worse, which is explained by the insufficient ability of the network to analyze smaller objects.

In [69], two experiments were conducted: one of them, on navigation at the district level (classification problem) and the other one, on navigation in the neighborhood (regression problem). The VGG16 architecture was used as a neural network [42], in which only two fully-connected layers were trained. Images of seven different district of Cairo were used for training. The navigation task at the district level was to correctly predict the district from UAV images. The network was able to distinguish between seven different districts with an average accuracy of 91.2%. For navigation in the neighborhood, search for the nearest neighbor was used, the input of which was the features of the 14th layer of the VGG16 neural network, while the feature map of the georeferenced reference image of the selected neighborhood was used to determine the absolute coordinates. The research results showed that the error in determining the coordinates was 200.75 m.

A new approach was proposed in [70], which uses a multi-stage multi-task CNN. A single image pass is sufficient to form a semantic segmented image and determine the location coordinates. The first stage of the proposed approach involved image segmentation aimed to detect roads. At the second stage, the decoder is implemented, which is capable of determining the location of the input image and calculating the longitude and latitude of the point. After the network was built, each of its outputs was trained independently in such a way that the loss function was a linear combination of binary cross-entropy and loss functions using the Dice coefficient [71]. The authors performed experiments on available datasets, such as Inria Aerial Image Labeling dataset (https://project.inria.fr/aerialimagelabeling/) and Massachusetts Buildings Dataset (https://www.kaggle.com/balraj98/massachusetts-buildings-dataset).

In the works considered below, a deep neural network is used to solve the problem of navigation in local areas of individual houses and quarters; both the UAV coordinates and orientation angles are determined. A quadrocopter was often used as a UAV, flying in an urban environment.

Thus, in [72], the PoseNet neural network was trained to determine six parameters of the quadrocopter position. The algorithm provided continuous estimation of location even when the scene included dynamic objects, such as people or cars, or had changing weather conditions. The PoseNet neural network was developed on the basis of the GoogLeNet neural network [73], in which the softmax layers were replaced by blocks for calculation of affine transformations, and one more fully-connected layer was added on the last block of affine transformations. Training and testing were carried out using the Cambridge Landmarks dataset of urban street objects (mi.eng.cam.ac.uk/projects/relocalisation/). Experiments have shown that the PoseNet neural network provides an accuracy of determining coordinates at a level of 2 m and orientation at a level of 3° over an area of 50000 m^2.

In [74], the authors used the SqueezePoseNet neural network [75], which is an adaptation of the SqueezeNet for more mobile platforms [76]. While SqueezeNet is intended to solve classification problems, SqueezePoseNet was modified to solve regression problems. Two fully-connected layers (position and rotation) were used to estimate the location. Leaky ReLU were used as activation functions [77], and layers of batch data normalization were used for higher learning rate. To increase the generalizing ability of the network, the data augmentation procedure was used. SqueezePoseNet was initialized with parameters trained on the ImageNet dataset [66]. The last fully-connected layers were additionally trained on the Shop Façade and Atrium datasets (Atrium dataset. https://www2.ipf.kit.edu/pcv2016/downloads/photos_atrium_recon-struction.zip). Experiments have shown that SqueezePoseNet provides a coordinate accuracy of 5.19 m and an orientation accuracy of 29.28°.

Conclusions

The algorithms presented in this group are used both for global positioning of an object over sufficiently large spaces (district, county, state) and for navigation in local areas. In the first case, only the coordinates of the object location are determined, while the accuracy estimation is often based not on errors in determining the coordinates, but on the probability of falling into a given area. In the second case, both coordinates and UAV orientation parameters are determined.

The considered algorithms cannot be used for spaces other than the original ones, since they require additional training procedures on data of the new location. Furthermore, such algorithms are more dependent on the amount of data used for training, which makes researchers create their own datasets that differ in resolution and size, and may also not take into account a number of factors, e.g. weather conditions.

Table 3. Comparative characteristics of image-matching navigation algorithms using neural networks

Work	Algorithms	Training data/Dimension w [m] × l [m]	UAV images/Dimension w [m] × l [m]	Mean error, vector state components
[67]	Xception	12 datasets/1500 × 1500 6.7 m/pix	1 dataset/1500 × 1500 6.7 m/pix	AED = 166.5 m (x, y)
[69]	VGG16	Google Maps/5251 × 9421, 3960 × 4874, 3960 × 2741	Simulated from Bing Maps/ 595 × 595 1.19 m/pix	AED = 200.75 m (x, y)
[70]	MSMT-Net	MBD, IAILD	MBD, IAILD	96.84% for 20 m tolerance (x, y)
[72]	PoseNet	Cambridge Landmarks[9]	Cambridge Landmarks[9]	AED = 2 m (x, y, z) and AED = 3° (ϕ, θ, ψ)
[74]	SqueezePoseNet	Shop Facade[9] and Atrium[10]	Shop Facade[9] and Atrium[10]	AED = 5.19 m (x, y, z) and AED = 29.28° (ϕ, θ, ψ)

Note AED – average Euclidian distance.

In general, such application of neural networks is being actively developed. The main challenge for scientists to overcome is to expand the area available for local navigation. The algorithms considered in this section cannot provide high navigation accuracy yet, but perhaps "new" network architectures will make it possible to reach new levels of accuracy.

The results of the analysis of the algorithms described above are presented in Table 3.

CONCLUSIONS

The article reviews the works on UAV navigation using vision systems, mainly since 2015. The works are grouped as follows:

— image-matching navigation algorithms using the correlation-extreme approach;

— image-matching navigation algorithms using key points;

— image-matching navigation algorithms using neural networks.

Algorithms based on image matching using correlation approaches often use segmentation neural networks for image preprocessing; however, little research has been done in this direction because of the low performance of such approaches. In general, these algorithms can be used for navigation in a certain type of terrain, usually with one-story residential buildings and dense road network, or at high altitudes involving hydrographic data. They can also be considered as corrective for the main navigation system.

Algorithms based on image matching using key points are the most widely used of all algorithms considered here. Detectors allow extracting the most "valuable" and "distortion-resistant" information from images and provide high reliability and accuracy in determining navigation parameters. The development of detectors is currently aimed at using neural networks to extract key points or form their more stable descriptors, allowing for accurate matching of the reference image and the image generated by the UAV. Algorithms of this group also use the optical flow when matching key points in order to reduce the area of their search in the image. Algorithms based on key point matching usually allow efficient estimation of the UAV coordinates and orientation (azimuth) and can be reliably used for navigation over extended spaces that do not have a large number of moving objects.

Algorithms using deep neural networks for image matching are currently the least studied and provide low accuracy. The use of such algorithms for UAV global positioning over sufficiently large spaces (district, neighborhood, state) makes it possible to estimate only the UAV coordinates, and when navigating in local areas, all 6 parameters of the object motion. The development of methods included in Group 3 aimed to minimize the size of networks in order to introduce them into UAV onboard computers and improve the generalizing properties for expanding the operation area of the UAV navigation systems seems to be the most promising.

REFERENCES

1. Candiago, S., Remondino, F., De Giglio, M., Dubbini, M., and Gattelli, M., Evaluating multispectral images and vegetation indices for precision farming applications from UAV images, *Remote Sensing*, 2015, vol. 7, no. 4, pp. 4026–4047. https://doi.org/10.3390/rs70404026
2. Akhloufi, M.A., Castro, N.A., and Couturier, A., *UAVs for wildland fires, Autonomous Systems: Sensors, Vehicles, Security, and the Internet of Everything*, 2018, vol. 10643. https://doi.org/10.1117/12.2304834
3. Akhloufi, M.A., Castro, N.A., and Couturier, A., Unmanned aerial systems for wildland and forest fires:

Sensing, perception, cooperation and assistance, *Drones*, 2021, vol. 5, no. 15.
https://doi.org/10.3390/drones5010015

4. Mokrova, M.I., Studying the effect of difficult fire conditions on the quality of observation and safety of UAV flight, *Izvestia YuFU. Tekhnicheskie nauki*, 2021.
https://doi.org/10.18522/2311-3103-2021-1-112-124
5. Jordan, S., Moore, J., Hovet, S., Box, J., Perry, J., Kirsche, K., Lewis, D., and Tse, Z. T. H., State-of-the-art technologies for UAV inspections, *IET Radar, Sonar & Navigation*, 2017, vol. 12, no. 2, pp. 151–164.
https://doi.org/10.1049/iet-rsn.2017.0251
6. Scherer, J., Yahyanejad, S., Hayat, S., Yanmaz, E., Andre, T., Khan, A., Vukadinovic, V., Bettstetter, C., Hellwagner, H., and Rinner, B., An autonomous multi-UAV system for search and rescue, *Proceedings of the First Workshop on Micro Aerial Vehicle Networks, Systems, and Applications for Civilian Use, DroNet '15, ACM, New York, USA*, 2015, pp. 33–38.
https://doi.org/10.1145/2750675.2750683.
7. Mittal, M., Mohan, R., Burgard, W., and Valada, A., Vision-based autonomous UAV navigation and landing for urban search and rescue, *Proceedings of the International Symposium on Robotics Research (ISRR)*, 2019.
https://doi.org/10.48550/arXiv.1906.01304.
8. Zoev, I.V., Markov, N.G., and Ryzhova, S.E., Intelligent system of computer vision of unmanned aerial vehicles for monitoring technological facilities of oil and gas companies, *Izvestiya Tomskogo politekhnicheskogo universiteta. Inzhiniring georesursov*, 2019, vol. 330, no. 11, pp. 34–49.
9. De Melo, C. F. E., E Silva, T. D., Boeira, F., Stocchero, J. M., Vinel, A., Asplund, M., and De Freitas, E. P., UAVouch: A secure identity and location validation scheme for UAV-networks, *IEEE Access*, 2021, vol. 9, pp. 82930–82946.
10. Mantelli, M., Pittol, D., Neuland, R., Ribacki, A., Maffei, R., Jorge, V., Prestes, E., and Kolberg, M., A novel measurement model based on abBRIEF for global localization of a UAV over satellite images, *Robotics and Autonomous Systems*, 2019, vol. 112, pp. 304–319.
https://doi.org/10.1016/j.robot.2018.12.006
11. Pazychev, D.B. and Mkrtchyan, V.I. Adaptive suboptimal Kalman filter in the SINS alignment problem. *Izvestiya Tul'skogo gosudarstvennogo universiteta. Tekhnicheskie nauki*, 2018, no. 5, pp. 60–73.
12. Conte, G. and Doherty, P., An integrated UAV navigation system based on aerial image matching, *Aerospace Conference*, 2008, pp. 1–10.
13. Viswanathan, A., Pires, B.R., and Huber, D., Vision-based robot localization across seasons and in remote locations, *International Conference on Robotics and Automation*, 2016, pp. 4815–4821.
14. Schmidt, G.T., GPS based navigation systems in difficult environments, *Gyroscopy and Navigation*, 2019, vol. 10, no. 2, pp. 41–53.
15. Peshekhonov, V.G., High-precision navigation independently of global navigation satellite systems data, *Gyroscopy and Navigation*, 2022, vol. 13, no. 1, pp. 1–6.
16. Sabatini, R., Moore, T., Hill, C., and Ramasamy, S., Avionics-based GNSS integrity augmentation performance in a jamming environment, *AIAC16: 16th Australian International Aerospace Congress, Engineers Australia*, 2015, pp. 469–479.
17. Groves, P.D., Jiang, Z., Rudi, M., and Strode, P., A portfolio approach to NLOS and multipath mitigation in dense urban areas, *Proceedings of the 26th International Technical Meeting of The Satellite Division of the Institute of Navigation, The Institute of Navigation*, 2013.
18. Zhuk, R.S., Zalesskii, V.A., and Trotskii, F.S., Visual navigation of an autonomously flying UAV with the aim of returning it to the starting point, *Informatika*, 2020, vol. 17, no. 2, pp. 17–24.
https://doi.org/10.37661/1816-0301-2020-17-2-17-24
19. Yuncheng, L., Zhucun, X., Gui-Song, X., and Liangpei, Z., A survey on vision-based UAV navigation, *Geospatial Information Science*, 21:1, 2018, pp. 21–32.
https://doi.org/10.1080/10095020.2017.1420509
20. Vizil'ter, Yu.V. and Zheltov, S.Yu. Using deep neural networks for data analysis, control and optimization in advanced aviation applications, In: *XII mul'tikonferentsiya po problemam upravleniya* (XII Multiconference on Control Problems (MKPU-2019)), 2019, Vol. 4, pp. 17–20.
21. Beloglazov, I.N. and Tarasenko, V.P., *Korrelyatsionno-ekstremal'nye sistemy* (Correlation-extreme systems), Moscow: Sov. radio, 1974.
22. Beloglazov, I.N. and Dzhandzhgava, G.I., *Osnovy navigatsii po geofizicheskim polyam* (Fundamentals of navigation by geophysical fields), Moscow: Nauka, 1985.
23. Stepanov, O.A. and Toropov, A.B., Nonlinear filtering for map-aided navigation, Part 2. Trends in the algorithm development, *Gyroscopy and Navigation*, 2016, vol.7, no.1, pp. 82–89.
24. Berkovich, S.B., Kotov, N.I., Lychagov, A.S., Panokin, N.V., Sadekov, R.N., and Sholokhov, A.V., Vision system as an aid to car navigation, *Gyroscopy and Navigation*, 2017, vol. 8, no. 3, pp. 200–208.
25. Sholokhov, A.V., Berkovich, S.B., Kotov, N.I., and Sadekov, R.N., Forming a trajectory of a map-matching navigation system by the criterion of minimum coordinate errors, *25th Anniversary Saint Petersburg International Conference on Integrated Navigation Systems,* 2018, pp. 1–3.
26. Stepanov, O.A. and Nosov, A.S., A Map-Aided Navigation Algorithm without Preprocessing of Field Measurements, *Gyroscopy and Navigation,* 2020, Vol. 11, No. 2, pp. 162–175.
27. Kozubovskii, S.F. *Korrelyatsionnye ekstremal'nye sistemy. Spravochnik*. (Correlation-extreme systems. Reference book), Kiev: Naukova Dumka, 1973.
28. Stepanov, O.A., *Metody otsenki potentsial'noi tochnosti v korrelyatsionno-ekstremal'nykh navigatsionnykh sistemakh* (Methods for Estimation of Potential Accuracy in Correlation-Extremal Navigation Systems), St. Petersburg, TsNII Elektropribor, 1993.
29. Poddar, Shashi et al., Evolution of visual odometry techniques. ArXiv abs/1804.11142 (2018): 12 p.
30. Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I.D., and Leonard, J.J., Simultaneous localization and mapping: Present, future, and the robust-perception age, *IEEE Transactions on Robotics* (cond. Accepted), 2016.

31. Ronneberger, O., Fischer, P., and Brox, T., U-Net: Convolutional networks for biomedical image segmentation, 2015. arXiv: 1505.04597.

32. Lowe, D.G., Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 2004, vol. 60, no. 2, pp. 91–110.

33. Isola, P., Zhu, J.-Y., Zhou, T., and Efros, A.A., Image-to-image translation with conditional adversarial networks, *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 1125–1134. https://doi.org/10.1109/CVPR.2017.632.

34. Mirza, M. and Osindero, S., Conditional generative adversarial nets, 2014. arXiv:1411.1784 [cs, stat].

35. Schleiss, M., Translating aerial images into street-map representations for visual self-localization of UAVs, *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2019, vol. 42, pp. 575–580. https://doi.org/10.5194/isprs-archives-XLII-2-W13-575-2019

36. Silva Filho, P., Shiguemori, E. H., and Saotome, O., UAV visual autolocalizaton based on automatic landmark recognition, *International Conference on Unmanned Aerial Vehicles in Geomatics*, 2017, pp. 89–94. https://doi.org/10.5194/isprs-annals-IV-2-W3-89-2017.

37. Alcantarilla, P.F., Nuevo, J., Bartoli, A., Fast explicit diffusion for accelerated features in nonlinear scale spaces, *British Machine Vision Conference (BMVC)*, 2013. https://doi.org/10.5244/C.27.13.

38. Ding, L., Zhou, J., Meng, L., and Long, Z., A practical cross-view image matching method between UAV and satellite for UAV-based geo-localization, *Remote Sensing*, 2021, vol. 13, no. 47. https://doi.org/10.3390/rs13010047

39. Zheng, Z., Wei, Y., and Yang, Y., University-1652: A multi-view multi-source benchmark for drone-based geo-localization, *Proceedings of the 28th ACM International Conference on Multimedia*, Seattle, WA, USA, October 2020, pp. 1395–1403.

40. Uchida, Yusuke. Local feature detectors, descriptors, and image representations: A survey. ArXiv abs/1607.08368 (2016).

41. Goforth, H. and Lucey, S., GPS-denied UAV localization using pre-existing satellite imagery, *International Conference on Robotics and Automation (ICRA). IEEE, Montreal, QC, Canada*, 2019, pp. 2974–2980. https://doi.org/10.1109/ICRA.2019.8793558.

42. Simonyan, K. and Zisserman, A., Very deep convolutional networks for large-scale image recognition, *Conference ICLR*, 2015. arXiv:1409.1556 [cs].

43. Fischler, M.A. and Bolles, R.C., Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of the ACM*, 1981, vol. 24, no. 6, pp. 381–395. https://doi.org/10.1145/358669.358692

44. Rublee, E., Rabaud, V., Konolige, K., and Bradski, G., ORB: An efficient alternative to SIFT or SURF, *2011 International Conference on Computer Vision*, 2011, pp. 2564–2571. https://doi.org/10.1109/ICCV.2011.6126544.

45. Seema, B.S., Hemanth, K., and Naidu, V.P.S., *Georegistration of aerial images using RANSAC algorithm, NCTAESD-2014. Vemana Institute of Technology, Bangalore*, 2014, pp. 1–5.

46. Saranya, K.C., Naidu, V.P.S., Singhal, V., and Tanuja, B.M., Application of vision-based techniques for UAV position estimation, *International Conference on Research Advances in Integrated Navigation Systems (RAINS)*, 2016, pp. 1–5. https://doi.org/10.1109/RAINS.2016.7764392.

47. Wang, X., Kealy, A., Li, W., Jelfs, B., Gilliam, C., May, S.L., and Moran, B., Toward autonomous UAV localization via aerial image registration, *Electronics*, 2021, vol. 10, no. 4. https://doi.org/10.3390/electronics10040435

48. Bay, H., Tuytelaars, T., and Van Gool, L., SURF: Speeded Up Robust Features, *A. Leonardis, H. Bischof, A. Pinz (Eds.), Computer Vision ECCV 2006, Lecture Notes in Computer Science, Springer Berlin Heidelberg*, 2006, pp. 404–417. https://doi.org/10.1007/11744023_32.

49. Stepanov, D.N., Application of Gabor wavelets in the problem of UAV navigation using a video camera, *Fundamental'nyye issledovaniya*, 2015, no. 12 (1), pp. 85–91.

50. Stepanov, D.N., Methods and algorithms for determining the position and orientation of an unmanned aerial vehicle using onboard video cameras, *Programmnye produkty i sistemy*, 2014, vol. 1, no. 1.

51. Nassar, A., Amer, K., ElHakim, R., and ElHelw, M., A deep CNN-based framework for enhanced aerial imagery registration with applications to UAV geolocalization, *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops*, 2018, pp. 1513–1523. https://doi.org/10.1109/ CVPRW.2018.00201.

52. Calonder, M., Lepetit, V., Strecha, C., and Fua, P., BRIEF: Binary Robust Independent Elementary Features, *K. Daniilidis, P. Maragos, N. Paragios (Eds.), Computer Vision ECCV 2010, Lecture Notes in Computer Science, Springer Berlin Heidelberg*, 2010, pp. 778–792. https://doi.org/10.1007/978-3-642-15561-1_56.

53. Masselli, A., Hanten, R., and Zell, A., Localization of unmanned aerial vehicles using terrain classification from aerial images, *Intelligent Autonomous Systems, Advances in Intelligent Systems and Computing, Springer International Publishing*, 2016, vol. 13, pp. 831–842. https://doi.org/10.1007/978-3-319-08338-4_60

54. Breiman, L., Random forests, *Machine Learning*, 2001, vol. 45, no. 1, pp. 5–32. https://doi.org/10.1023/A:1010933404324

55. Bianchi, M. and Barfoot, T.D., UAV localization using autoencoded satellite images, *IEEE Robotics and Automation Letters*, 2021, vol. 6, no. 2, pp. 1761–1768.

56. Shan, M., Wang, F., Lin, F., Gao, Z., Tang, Y. Z., and Chen, B. M., Google map aided visual navigation for UAVs in GPS-denied environment, *IEEE International Conference on Robotics and Biomimetics (ROBIO)*, 2015, pp. 114–119. https://doi.org/10.1109/ROBIO.2015.7418753.

57. Dalal, N. and Triggs, B., Histograms of oriented gradients for human detection, *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*

(CVPR'05). San Diego, CA, USA, 2005, vol. 1, pp. 886–893.
https://doi.org/10.1109/CVPR.2005.177.

58. Horn, B.K.P. and Schunck, B.G., Determining optical flow, *Artificial Intelligence*, 1981, vol. 17, no. 1, pp. 185–203.
https://doi.org/10.1016/0004-3702(81)90024-2
59. Mughal, M.H., Khokhar, M.J., and Shahzad, M., Assisting UAV localization via deep contextual image matching, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, vol. 14, pp. 2445–2457.
60. He, K., Zhang, X., Ren, S., and Sun, J., Deep residual learning for image recognition, 2015. arXiv:1512.03385v1 [cs.CV]
61. Rocco, I., Cimpoi, M., Arandjelovic, R., Torii, A., Pajdla, T., and Sivic, J., Neighbourhood consensus networks, *32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montréal, Canada,* 2018.
62. Everingham, M., Gool, L. V., Williams, C. K. I., Winn, J., and Zisserman, A., The PASCAL visual object classes (VOC) challenge, *International Journal of Computer Vision*, 2010, vol. 88, pp. 303–338.
63. Taira, H., Okutomi, M., Sattler, T., Cimpoi, M., Pollefeys, M., Sivic, J., Pajdla, T., and Torii, A., InLoc: Indoor visual localization with dense matching and view synthesis, 2018. arXiv:1803.10368 [cs]
64. Tanchenko, A.P., Fedulin A.M., Bikmaev R.R., and Sadekov R.N., UAV navigation system autonomous correction algorithm based on road and river network recognition, *Gyroscopy and Navigation*, 2020, vol. 11, no. 4, pp. 293-299.
https://doi.org/10.1134/S2075108720040100
65. Krizhevsky, A., Sutskever, I., and Hinton, G. E., Imagenet classification with deep convolutional neural networks, *Communications of the ACM*, 2017, vol. 60, no. 6, pp. 84–90.
https://doi.org/10.1145/3065386
66. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., et al. ImageNet Large Scale Visual Recognition Challenge, *International Journal of Computer Vision (IJCV)*, 2015, vol. 115, no. 3, pp. 211–252.
https://doi.org/10.1007/s11263-015-0816-y
67. Harvey, W., Rainwater, C., and Cothren, J., Direct Aerial Visual Geolocalization Using Deep Neural Networks, *Remote Sensing*, 2021, vol. 13,
https://doi.org/10.3390/rs13194017
68. Chollet, F., Xception: Deep Learning with Depthwise Separable Convolutions, 2017. arXiv:1610.02357 [cs]
69. Amer, K., Samy, M., ElHakim, R., Shaker, M., and El-Helw, M., Convolutional Neural Network-Based Deep Urban Signatures with Application to Drone Localization, *IEEE International Conference on Computer Vision Workshops (ICCVW)*, 2017, pp. 2138–2145.
https://doi.org/10.1109/ICCVW.2017.
70. Marcu, A., Costea, D., Slusanschi, E., and Leordeanu, M., A Multi-Stage Multi-Task Neural Network for Aerial Scene Interpretation and Geolocalization, 2018. arXiv:1804.01322 [cs].
71. Jadon, S., A survey of loss functions for semantic segmentation, *IEEE International Conference on Computational Intelligence in Bioinformatics and Computational Biology*, 2020.
https://doi.org/10.1109/CIBCB48159.2020.9277638.
72. Kendall, A., Grimes, M., and Cipolla, R., Posenet: A convolutional network for real-time 6-dof camera relocalization, *Proceedings of the IEEE International Conference on Computer Vision, Santiago, Chile*, 2015, pp. 2938–2946.
73. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A., Going deeper with convolutions, *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA*, 2015, P. 1–9.
74. Mueller, M.S. and Jutzi, B., UAS Navigation with SqueezePoseNet—Accuracy boosting for pose regression by data augmentation, *Drones*, 2018, vol. 2, no. 7.
https://doi.org/10.3390/drones2010007
75. Mueller, M.S., Urban, S., and Jutzi B., SqueezePoseNet: Image based pose regression with small convolutional neural networks for real time UAS navigation, *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci*, 2017, vol. 4, pp. 49–57.
76. Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., and Keutzer, K., SqueezeNet: AlexNet-level accuracy with 50× fewer parameters and <0.5 mb model size, 2016. arXiv:1602.07360.
77. Maas, A.L., Hannun, A.Y., and Ng, A.Y., Rectifier nonlinearities improve neural network acoustic models, *Proceedings of the International Conference on Machine Learning, Atlanta, GA, USA*, 2013, vol. 30.

SPELL: 1. OK