GreedyCenters:
Satellite imagery adaptive sampling method for artificial neural networks training

Oleg Gvozdev[1*]

[1] Moscow State University of Geodesy and Cartography

Abstract. The one of many significant particularities of satellite imagery is large size of images within orders of magnitude exceeds capability of modern GPGPU to train neural networks on its full size. On the other hand satellite imagery tends to be limitedly available. Moreover, the objects of interest tends to constitute a small fraction of whole dataset. This leads to the demand of sample extraction and augmentation method specialized on satellite imagery. Yet this area is immensely underrated so almost all widely used method limited to grid-based sample extraction and augmentation via combinations of 90-degrees rotations and mirroring on vertical or horizontal axes. This paper proposes the domain-agnostic method of sample extraction and augmentation. Adoption of this method to specific subject area is based on domain-specific way to generate significance field of image. In contrast to trivial greedy solutions and more advanced stochastic optimization methods the design of proposed method is focused on maximizing per-step progress. This makes its performance reasonably good even without low-level optimizations without significant quality loss. It can be easily implemented using widely known and open source software libraries.

1 Introduction

The design and implementation of data preparation pipeline is essential part of building any machine learning model. Gaining the best results relies not just on deep understanding of particularities and nuances of subject area, involved data sources and models yet on exploiting them ingenuently and efficiently.

Processing satellite imagery is one of such distinct areas. It very different from ordinary visual experience and photographic technologies. The satellite images can easily exceed resolution of 30000x30000 pixels, contain arbitrary number of spectral channels of different spatial resolution. They can contain void areas, that have no data at all. The shown territories can be highly variable in terms of relevant information spatial density. The objects of interest usually quantitatively and arealy unbalanced with its environment within an orders of magnitude.

The deep analysis and explanation of all mentioned particularities are described in [1].

---

* Corresponding author: gvozdev@miigaik.ru

For the purpose of this research most important distinction of satellite imagery is its large non uniform area, which cannot be processed by modern GPGPU during neural network training as a single raster, yet cannot be downsampled without loss of important features.

This leads to the demand of sample extraction method specialized on satellite imagery. Which is being coupled with image augmentation methods is the core of dataset preparation pipeline.

Sadly, this area is not as developed as much in the corresponding tasks of processing ordinary visual experience images. For example it's a lot of software libraries that provides different augmentation techniques for them [2, 3]. Yet researches on sampling and augmentation of satellite imagery is almost limited to manually extracted samples augmented by combinations of 90-degrees rotations and mirroring on vertical or horizontal axes [4-6]. Sometimes automatic sample extraction based on uniform grid sampling is used. The most of papers in this field are not mention any image sampling or augmentation techniques at all. This approach cannot be considered as general case solution.

The first attempt to propose generic approach to this problem done by author is given in [1]. Its general idea is decomposition of the problem via extracting two customizable domain-specific parts:

1. Obtaining of image significance field.
2. Performing of image fragment augmentation.

While the rest part of the solution stays domain-agnostic. This part tries to find a best coverage of provided significance field with quadrangle fragments according to values of field. In following each of the found fragments are extracted, transformed and augmented independently.

This problem is a special case of known NP-complete set cover problem [7]. In the original work [1] its said that even naive greedy algorithm solves task in a such setting good enough for practical applications. This implemented by generation of large amount of proposition quadrangles. Followed by iterative choice the one that covers current residual significance field the best and adding it to resulting subset. The process repeated until required coverage is not reached.

The practice with this solution show its ability to solve class imbalance issues and non-uniformity of information spatial density and provide solid base for data preparation pipeline.

Yet there are some drawbacks found.

The greed algorithm of fragments generation relays on recomputation of residual significance field after each step generating only one fragment. This is extremely computationally intensive and makes scheme not applicable for streaming sample generation. Especially while images tends to become larger. Optimization of residual significance field recomputation is influential yet inherently limited approach. So it not considered in this research.

Moreover, with growth of image resolution relative size of sample fragment become insignificant. This makes direct optimization of all quadrangle parameters (coordinates of four points) worthless.

This leads author to a more simple yet more reliable and performant approach, preserves the proven advantages of the original scheme, especially the concept of significance field. The method based on it is described in this paper.

2 Methods

The core issue of method is finding a set of quadrangles that covers significance field according to it values. The exact position of all quadrangle points, its rotation and flip is

insignificant, as is the exact coverage of each of them. Consequently, the process can be split in two parts:

1. Choosing the set of points, which covers given significance field.
2. Generate random quadrangles with centers in those points.

The coverage of each point can be estimated by convolution on significance field with window function represent coverage of central part of coming quadrangle.

While it's obviously a single global maxima on each step of optimization, there can be several local maxima neighbourhood areas, which can be processed independently. That way makes it possible to produce multiple points on each step.

The processing of each local maxima neighbourhood should be variable part of method.

Thereby deterministic implementation allows to get reproducible and describable results. While stochastic implementation can be utilized for generation of new samples each time. Or for selection best result from several runs of procedure.

Using the concepts described above author designed the method which scheme shown on fig. 1.
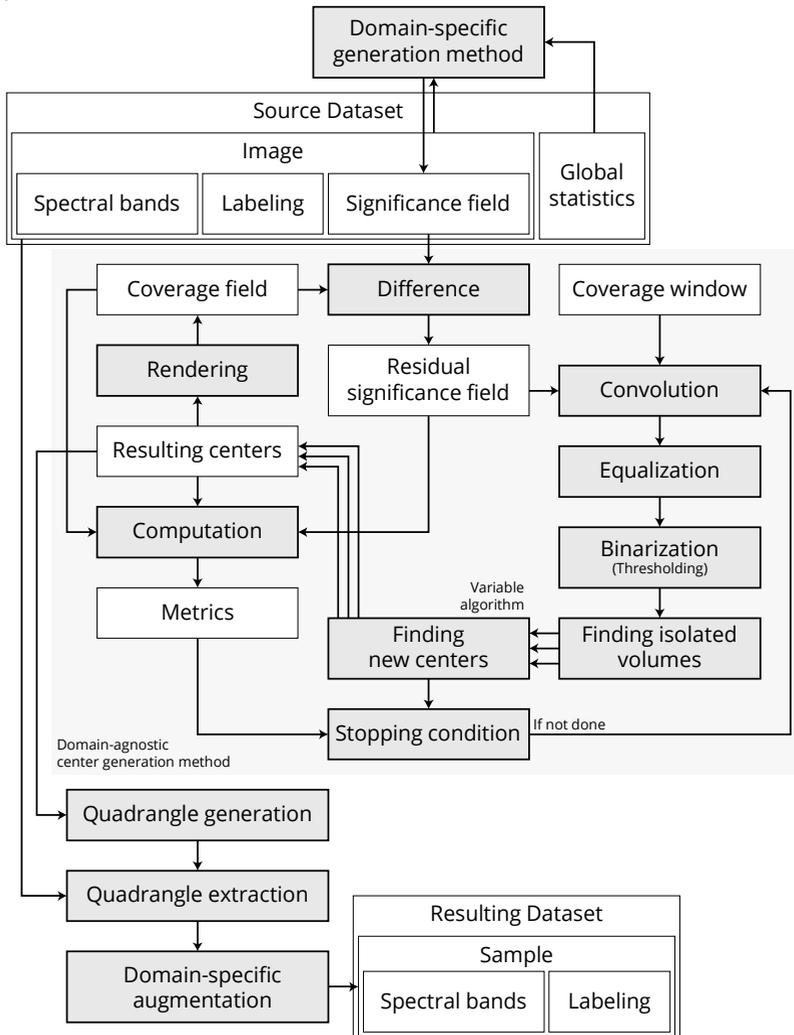

Fig. 1. The scheme of proposed method for satellite imagery sample extraction.

2.1 Implementation details

The exact iteration body can be implemented as follows:

1. Acquire binearized mask for current residual significance field.

2. Find isolated volumes in mask via composition of scipy.ndimage.label and scipy.ndimage.find_objects functions from SciPy [8].

3. Apply center choice algorithm for each found isolated volume.

4. Add found centers to the resulting array.

5. Update coverage field via deconvolution of centers according to the chosen window function.

6. Update residual significance field via subtracting coverage field from source significance field.

7. Update metrics.

Binearized mask acquisition in turn assumes three steps:

1. Convolution of residual significance field according to the chosen window function.

2. Significance field histogram equalization.

3. Threshold filtration by value 0.9.

The convolution of large images according to Convolution Theorem [9] can be optimized via exploiting its FFT representation. This implemented via scipy.signal.convolve from SciPy [8].

The histogram equalization method can itself be one of widely available [10]. The exact implementation exploits trivial min-max normalization.

The center choice algorithm can be one of follows:

1. Deterministically choose global maxima of significance field in isolated volume.

2. Randomly choose any point in isolated volume considering all points as equal.

3. Randomly choose any point in isolated volume using significance field value as weight.

The followed extraction and transformation procedures based on resampling procedure skimage.transform.warp implemented in SciKit-Image [11].

For the purposes of methods benchmark there was additionally implemented greedy global maxima version of center chooser.

2.2 Metrics

Denote the following:

1. Significance field – S.

2. Coverage field – C.

Define the following:

1. Overcoverage field $C_+ = max(C - S, 0)$.

2. Undercoverage field $C_- = max(S - C, 0)$.

3. Target coverage $T = \sum S$

The following metric are used during methods benchmark:

1. Undercoverage ratio $R_- = \frac{\sum C_-}{T}$.

2. Overcoverage ratio $R_+ = \frac{\sum C_+}{T}$.

3. Output coverage ratio as $R_c = 1 - R_-$.

The finishing condition of method is based on thresholding of Output coverage ratio.

3 Results

For the relative benchmark of proposed method variances the simple test image was chosen (fig. 2). The significance field was generated involving its local entropy [12] and object classes (fig. 2).
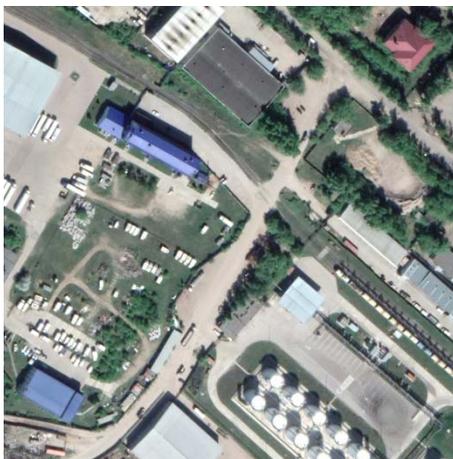
Fig. 2. Benchmark image. 1000x1000 px.
Source: Google Earth, CNES / Airbus, Maxar Technologies.
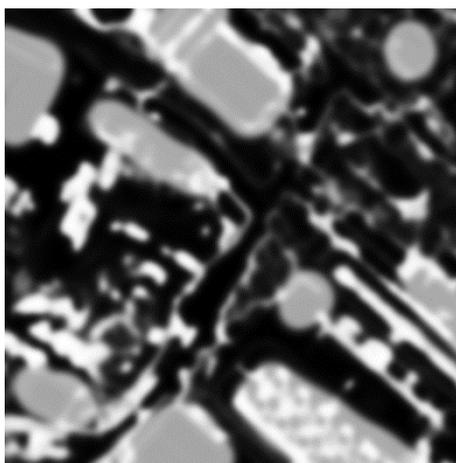
Fig. 3. Significance field for benchmark image.

As a window function circular linear falloff of different sizes was chosen: 128x128, 96x96, 64x64, 32x32. This allows to evince the changes of performance proportionally to decrease of relative size of desirable fragments.

The resulting residual significance fields obtained via different algorithms and window functions is show on fig. 4-7. Undercoverage shown in reds. Overcoverage shown in blues.

The detailed statistics is given in table 1.

Table 1. Resulting statistics of proposed method benchmark.

Method variant	Sample size	Steps	Centers	R_-	R_+
Globally greed method	128x128	60	60	0.0267	1.2946
	96x96	98	98	0.0164	1.2356
	64x64	198	198	0.0094	1.1506
	32x32	606	606	0.0099	0.8318
Locality-based greed method	128x128	19	42	0.0748	0.8117
	96x96	26	71	0.0621	0.7980
	64x64	37	154	0.0370	0.7756
	32x32	52	556	0.0197	0.6977
Locality-based random method	128x128	19	41	0.0745	0.8193
	96x96	26	74	0.0574	0.8527
	64x64	33	148	0.0440	0.7202
	32x32	53	563	0.0196	0.7206
Locality-based weighted random method	128x128	20	43	0.0628	0.8887
	96x96	25	71	0.0694	0.7905
	64x64	31	150	0.0402	0.7508
	32x32	57	560	0.0206	0.7083

128x128

96x96

64x64

32x32
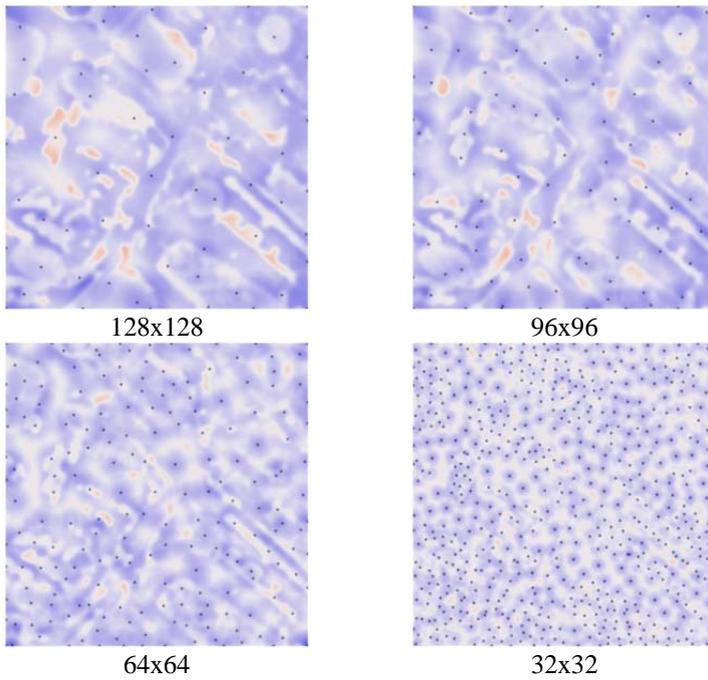
Fig. 4. The resulting residual significance fields obtained
via globally greed variant.

128x128

96x96

64x64

32x32
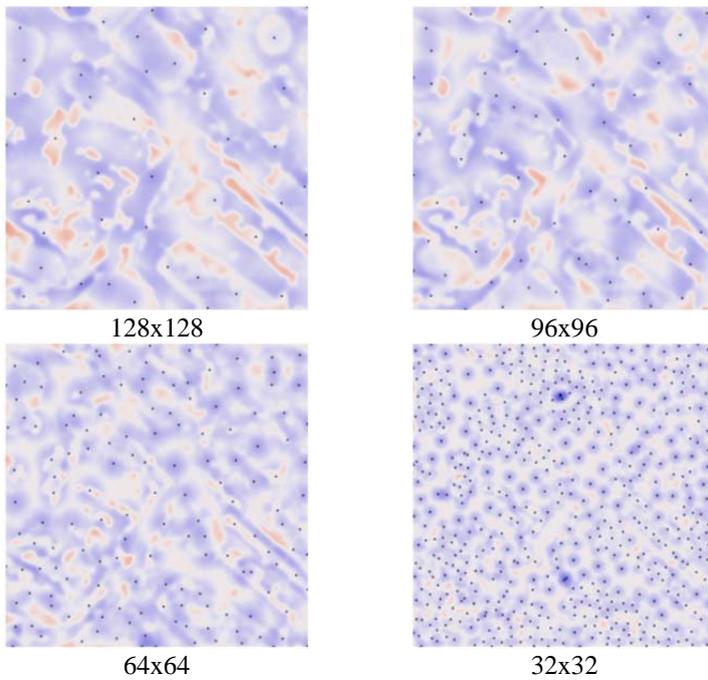
Fig. 5. The resulting residual significance fields obtained
via locality-based greed algorithm.

128x128

96x96

64x64

32x32
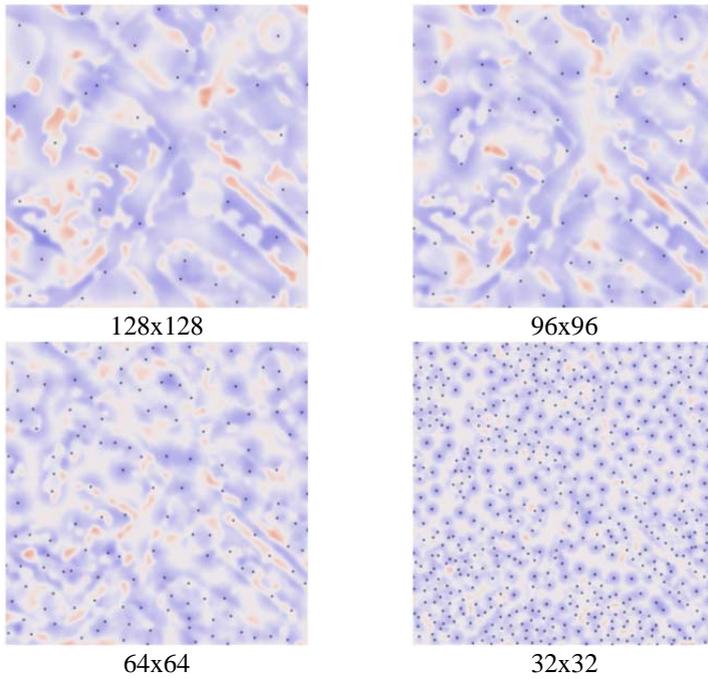
Fig. 6. The resulting residual significance fields obtained
via locality-based random algorithm.

128x128

96x96

64x64

32x32
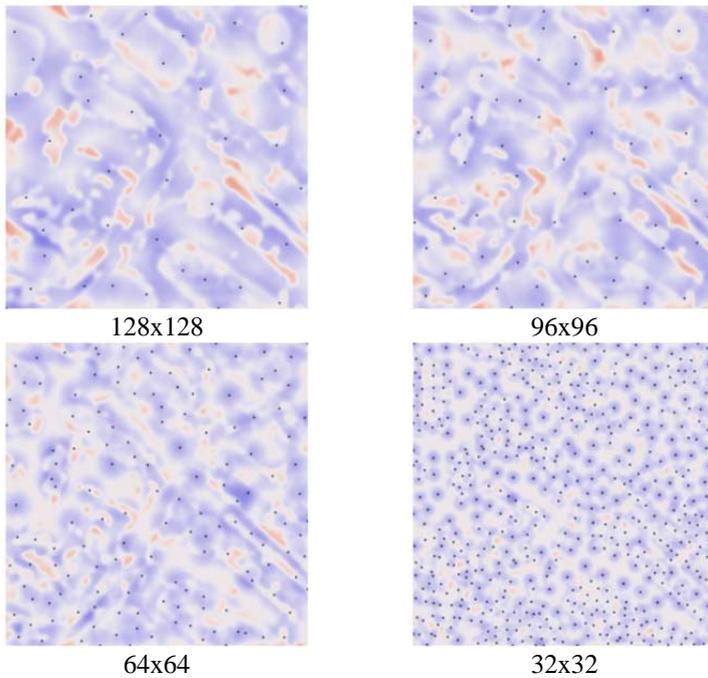
Fig. 7. The resulting residual significance fields obtained
via locality-based weighted random algorithm.

The dynamic of convergence of method variances on different sample sizes show on fig. 8.
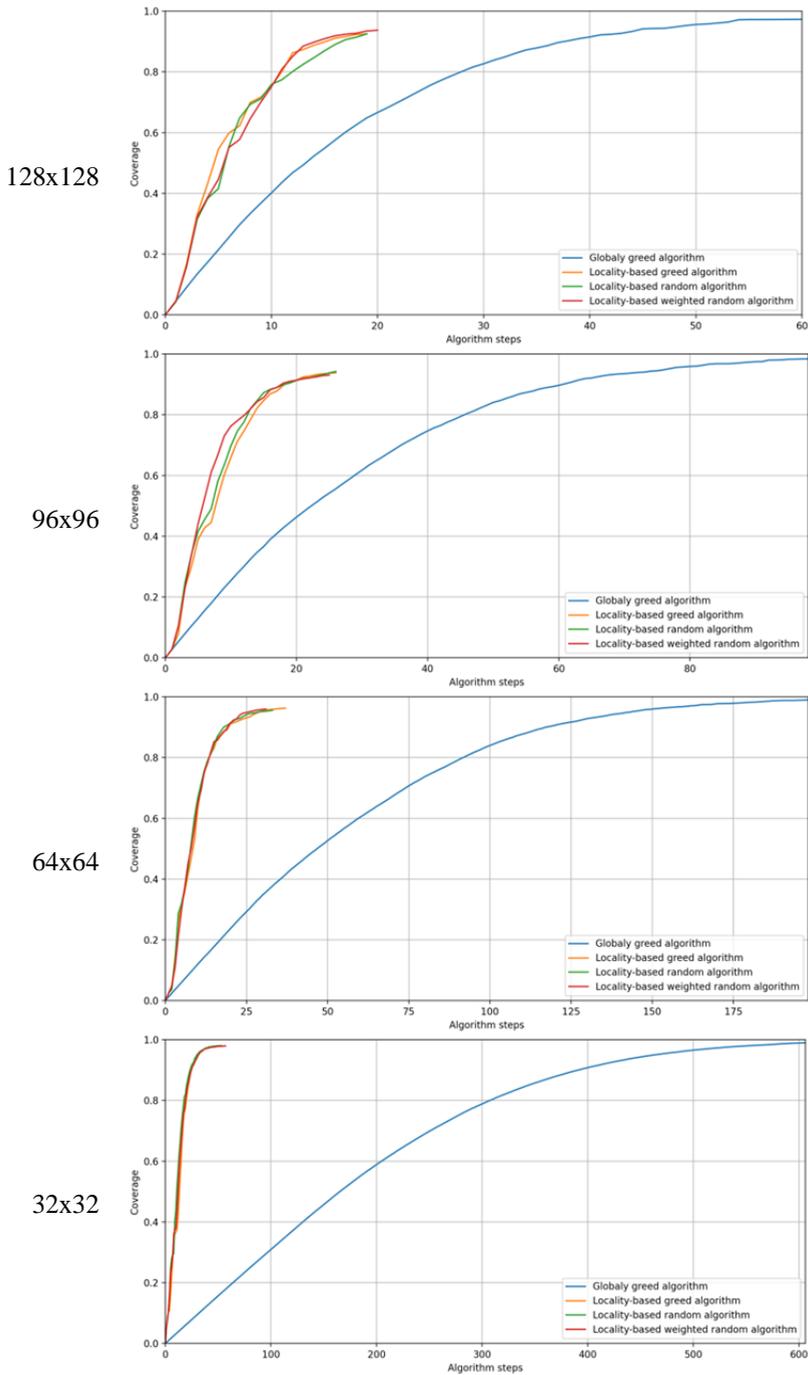

Fig. 8. The resulting residual significance fields obtained via locality-based weighted random algorithm.

4 Discussion

The gained statistic shows that while having comparable quality all locality-based algorithms converges much faster respectively to decrease relative size of sampling fragment size.

Visual analysis of resulting residual significance fields and found center sets shows that difference in that particular example is negligible. As the difference between locality-based algorithms.

The exact influence of algorithm parameters and exact significance field on gained sampling results and followed artificial neural network models training is open topic and requires as experimental and theoretical work.

The author also researched other stochastic methods such as simulated annealing and genetic algorithms. Their results are comparable to greedy algorithms in this particular task yet with at least an order of magnitude more computations. So it can be concluded that they are not practical for this task. At least in given setting.

There are still a lot of improvement possibilities such as automatic hyperparameter tuning, utilization of GPGPU or parallel computation via methods like work-stealing [13].

The most important prospective research is applying developed method in a streaming scenario for generating training samples based on significance field built based on feedback from current accuracy metrics of model on different parts of source images.

5 Conclusion

It known that artificial neural network can show state-of-the-art results in image processing task. But it requires high quality large representative training datasets.

There are distinct subject areas that has inheritant limitation of data availability. Processing satellite imagery is one of those.

The deficiency of training data comes from many reasons. And the augmentation of data is proven way to partially deal with that deficiency. Yet there are only a few simple solutions are widely known. This indicates that mentioned issues are immensely underrated.

This paper proposes method that gives a researchers and developers the simple and productive approach to that issue. The proposed method can be adapted to virtually any subject area just by customizing significance field generation. It can be easily implemented using widely known and open source software libraries.

References

1. O.G. Gvozdev, I.B. Abbasov, Recognition and Perception of Images, Fundamentals and Applications, 369-444 (2021).
2. A. Buslaev , V.I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, A.A. Kalinin, *Albumentations: Fast and Flexible Image Augmentations*, Information, **11**, 125 (2020).
3. Imgaug, python library, URL: https://github.com/aleju/imgaug
4. A. Kozionov, K. Tsyganov, J. Bologov, A. Andreev, I. Gorokhov, *Automated Area Assessment of Objects Using Deep Learning Approach and Satellite Imagery Data*, The 6th International Conference on Analysis of Images, Social Networks, and Texts (2012).
5. F. Waldner, F. Diakogiannis, *Deep learning on edge: extracting field boundaries from satellite images with a convolutional neural network* (2019).

6. Landuse Classification from Satellite Imagery using Deep Learning, URL: https://www.slideshare.net/Hadoop_Summit/landuse-classification-from-satellite-imagery-using-deep-learning

7. V. V. Vazirani, Approximation Algorithms, 396 (2001).

8. P. Virtanen, R. Gommers, T.E. Oliphant, et al., *SciPy 1.0: fundamental algorithms for scientific computing in Python*. Nat Methods, **17**, 261–272 (2020).

9. R. Bracewell, *Convolution Theorem*, *The Fourier Transform and Its Applications*, 3rd ed. New York: McGraw-Hill, 108-112 (1999).

10. Scikit-image, Local Histogram Equalization, URL: https://scikit-image.org/docs/dev/auto_examples/color_exposure/plot_local_equalize.html

11. S. Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, T. Yu, *scikit-image: Image processing in Python*. PeerJ, **2**, e453 (2014).

12. Scikit-image, Entropy, URL: https://scikit-image.org/docs/dev/auto_examples/filters/plot_entropy.html

13. R. D. Blumofe, C. E. Leiserson, *Scheduling multithreaded computations by work stealing*, J ACM, **46**, 720–748 (1999).