CV for anomaly detection in industrial applications

Viacheslav Kozitsin, Iurii Katser, Arman Alahyari, Anton Hinneck, Rahim Tariverdi

Skolkovo Institute of Science and Technology

Abstract

Keywords: deep learning, computer vision, anomaly detection, fault detection, technical systems

---

1. INTRODUCTION

Anomaly detection problems have a great importance in industrial applications, because anomalies usually represent faults, failures or the emergence of such. To detect these automatically, we propose deep learning algorithms for anomaly / fault detection and their classification. We develop an algorithm achieving Segmentation Defects classification Its performance will be examined on two different problems: anomaly detection in oil pipelines and fault detection on transmission system components. Both these systems can span over thousands of kilometers, which makes manual inspection very costly. How such computer vision techniques can be applied on an industrial scale is discussed in the following.

The damage of pipelines that transport petroleum and gas products lead to serious environmental problems. Eliminating breakthroughs and their consequences is expensive. To avoid accidents, it is recommended to improve diagnostics quality and to increase the frequency of pipeline inspectors deployment. These are specialized robots (inspectors) that Identify defects and evaluate the thickness of the pipeline. Most of the techniques currently used for analysis relate to heuristical or traditional ML approaches and don't involve CNN or DL algorithms.

Transmission infrastructure, on the other hand, physically connects power sources and consumers extending over thousands of kilometers. Many different components exist, which is why maintaining power grids is a serious cost factor for transmission system operators (TSOs). Automated fault detection could potentially help to decrease costs. To use an automated visual system however, specific infrastructure first needs to be identified and segmented, to then perform any type of fault detection. Therefore, the proposed deep learning-based approach, to segment and identify faulty components (i.e. insulators) in handheld footage, will be tested regarding its viability. If a deep-learning based approach would prove reliable, drones could monitor equipment automatically.

2. LITERATURE REVIEW

2.1. Pipelines defects diagnostics

Magnetic Flux Leakage (MFL) technique is the most common approach for oil and gas pipelines nondestructive diagnostics. The data obtained in the pipeline inspection process is primarily analyzed by traditional machine learning (ML) methods. A comparison of performance among different ML methods for defects identification problem is presented in [1]. The main challenge for this approach is creating informative and important features that will be used as an input for ML methods. Usually, these diagnostics features are generated using expert knowledge and manually-created heuristics. It imposes the limitation on defects detection problem solving quality. A variety of most successful features is presented and analyzed in details in [2].

To automate the process of feature generation Convolutional Neural Networks (CNNs) can be used. As an advantage, they can solve the defects detection task and at the same time. In literature there are samples of applying CNNs for defects detection [3], welds defect detection [4], welds detection and defects detection and classification [5], defect size estimation [6]. For all mentioned applications, CNNs outperformed existing traditional approaches. Since DL relatively recently showed great progress in their results achievements, there are just a few works dedicated to MFL data analysis using DL. For instance, we could not find any works on applying CNNs to defects segmentation task, despite the importance of this problem solving according to [3]. In this work, we want to solve two different tasks:

1. Defects detection (Picture classification task).
2. Defects segmentation (Instance segmentation task).

2.2. Unet

asd [7]

3. CONTRIBUTION

Viacheslav Kozitsin.

Iurii Katser.

Arman Alahyari.

Anton Hinneck.

Rahim Tariverdi.

4. 3RD PARTY CODE LIST

5. METHODS PIPELINES

5.1. Raw data

Example of the MFL data is shown in Fig. 1.

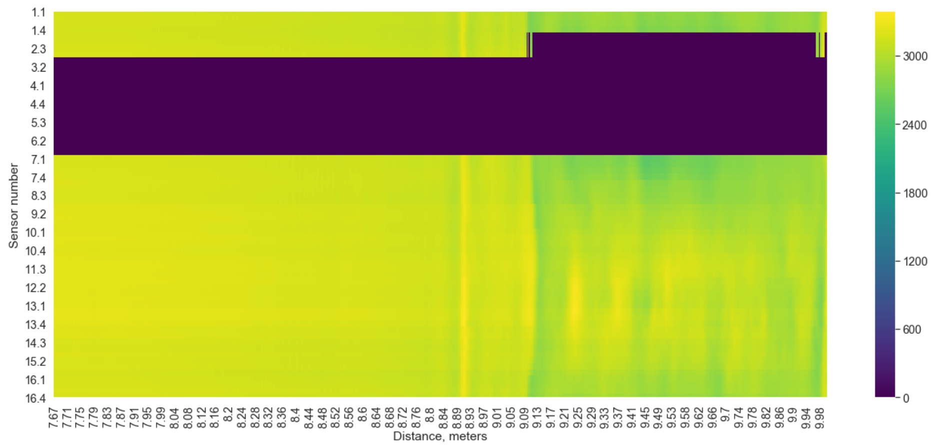

Figure 1: Example of the MFL data

Each MFL dataset provides information about single inspection tool run. Selected for the research dataset represents 15162.85 meters length pipeline part. It has 745 defects of different types and 1462 welds. Fig. 2 shows examples of normal data, data with the weld and defect. Attached to the dataset technical report contains information about welds and defects location, defects types and sizes and other related data.

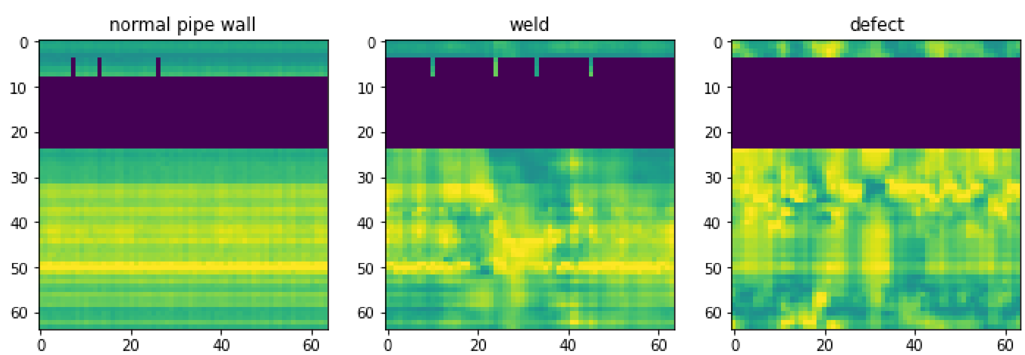

Figure 2: Example of the MFL data

Raw data has several issues that don't allow to solve CV problems without proper preprocessing. They are:

1. Sensors malfunctions (zeroed values cause bold horizontal line in Fig. 1);
2. Displaced origins between data and report files coordinates;
3. Inaccurate annotations, e.g. missed defect, wrong defect location, etc.

In addition to preprocessing, data were annotated for the segmentation task.

5.1.1. Sensors malfunctions problem

To deal with sensors malfunctions we suppose to fill the gaps (zeroed values) with values calculated by different methods:

1. Scaling of picture values to [0.5 : 1] range. Abnormal values (< 2000) are equal to 0.
2. Abnormal values are equal to the mean of normal values from one picture.
3. Abnormal values are equal to the mean of normal values over the column.
4. Abnormal values are equal to the mean of neighboring sensors over the column.
5. Abnormal values are equal to the interpolation results over the column.
6. Replacing Abnormal values to NaN and rerange in accordance skimage procedure
7. Rerange initial set of values to 0...255 uint8 range.
8. Rerange initial set of values to 0...1 float range.

The results of all applied methods are presented in Fig. 3.

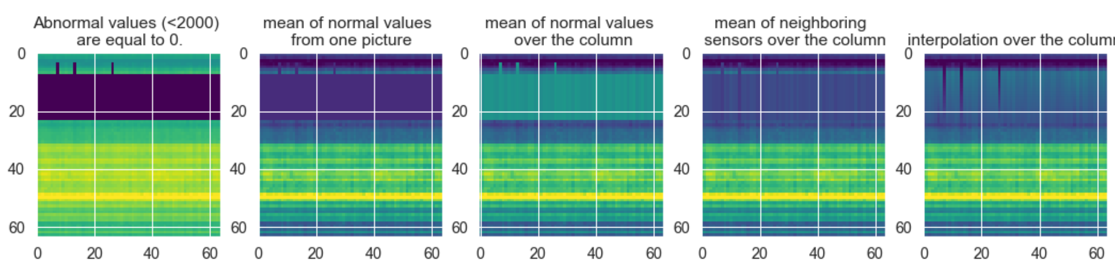

Figure 3: The results of missing values filling methods

5.1.2. Displaced origins problem

Since the data on the location of the robot did not match the defect location data from the report, it was necessary to merge the data. The key factor here turned out to be that the values of signal from magnetic flux sensors grow at the weld site, see picture 4.

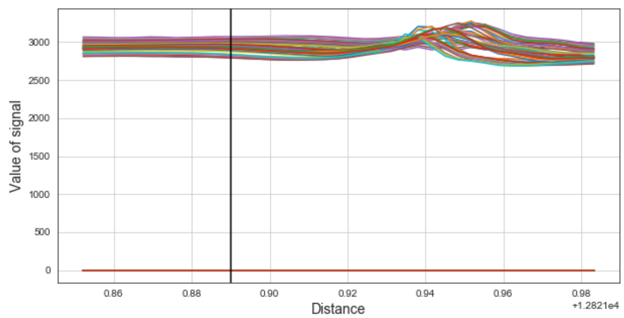

Figure 4: Location of weld. Black vertical line in accordance with defect location data from the report. Others lines are values from sensors.

The locations of maxima of these values have been founded and then combined with welds.

5.1.3. Inaccurate annotations problem

5.2. Augmentation

Although we have a lot of data, we don't have a lot of defects and welds in comparison with normal pipe wall instances. We use augmentation procedure to balance classes of pictures and increase model's quality by increasing number of instances in small classes (defects, welds). As an augmentation tool we use Albumentations library [8]. For welds pictures we apply following augmentations:

1. Rotate (limit=180, p=1),
2. VerticalFlip (p=1),
3. HorizontalFlip (p=1),
4. ElasticTransform (p=1, alpha=20, sigma=120 * 0.05, alpha affine=120 * 0.03),
5. GridDistortion (p=1),
6. OpticalDistortion (p=1, distort limit=2, shift limit=0.5),

And for defects we apply following ones:

1. Rotate (limit=180, p=1),
2. VerticalFlip (p=1),
3. HorizontalFlip (p=1),
4. ElasticTransform (p=1, alpha=20, sigma=120 * 0.05, alpha affine=120 * 0.03),
5. GridDistortion (p=1),
6. OpticalDistortion (p=1, distort limit=2, shift limit=0.5),

Examples of augmentations are shown in Fig. 5.

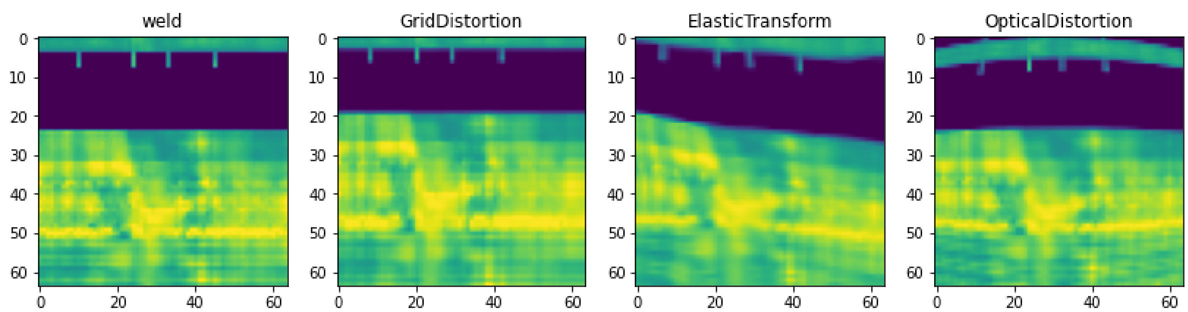

Figure 5: Example of the MFL data

Table 1: Dataset size for pipeline defects detection problem

Class name	train	test
Before augmentation		
Normal pipe wall	11106	584
Defect	1130	282
Weld	569	142
After augmentation		
Normal pipe wall	11106	584
Defect	8897	282
Weld	7680	142

6. RESULTS PIPELINES

7. RESULTS PIPELINES

8. METHODS TRANSMISSION

9. RESULTS TRANSMISSION

10. CONCLUSION

References

[1] A. Khodayari-Rostamabad, J. Reilly, N. Nikolova, J. Hare, S. Pasha, Machine learning techniques for the analysis of magnetic flux leakage images in pipeline inspection, IEEE Transactions on Magnetics 45 (8) (2009) 3073–3084. doi:10.1109/tmag.2009.2020160.

[2] S. D.A., Development of the magnetic method of non-destructive testing due to automation of data processing and optimization of defect detection algorithms, Ph.D. thesis, National Research University "Moscow Power Engineering Institute" (2017).

[3] J. Feng, F. Li, S. Lu, J. Liu, D. Ma, Injurious or noninjurious defect identification from MFL images in pipeline inspection using convolutional neural network, IEEE Transactions on Instrumentation and Measurement 66 (7) (2017) 1883–1892. doi:10.1109/tim.2017.2673024.

[4] Oil pipeline weld defect identification system based on convolutional neural network, KSII Transactions on Internet and Information Systems 14 (3) (mar 2020). doi:10.3837/tiis.2020.03.010.

[5] L. Yang, Z. Wang, S. Gao, Pipeline magnetic flux leakage image detection algorithm based on multiscale SSD network, IEEE Transactions on Industrial Informatics 16 (1) (2020) 501–509. doi:10.1109/tii.2019.2926283.

[6] S. Lu, J. Feng, H. Zhang, J. Liu, Z. Wu, An estimation method of defect size from MFL image using visual transformation convolutional neural network, IEEE Transactions on Industrial Informatics 15 (1) (2019) 213–224. doi:10.1109/tii.2018.2828811.

[7] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Lecture Notes in Computer Science, Springer International Publishing, 2015, pp. 234–241. doi:10.1007/978 $-$ $3-319-24574-4_28$.

[8] A. Buslaev, V. I. Iglovikov, E. Khvedchenya, A. Parinov, M. Druzhinin, A. A. Kalinin, Albumentations: fast and flexible image augmentations, Information 11 (2) (2020) 125.