Machine Learning aided noise filtration and signal classification for CREDO experiment

Łukasz Bibrzycki *et al.*
on behalf of CREDO collaboration

Pedagogical University of Krakow

lukasz.bibrzycki@up.krakow.pl

12-23 July 2021

CREDO
THE QUEST FOR THE UNEXPECTED

Overview

What is CREDO ?

The **C**osmic **R**ay **E**xtremely **D**istributed **O**bservatory (CREDO) experiment is designed for the large-scale study of various radiation of extraterrestrial origin, collectively known as cosmic rays.

It utilizes the CREDO Detector mobile application.

The project follows the citizen science philosophy and engages various groups of people, mostly non-scientists.

By installing CREDO Detector App they turn their smartphones into mobile cosmic ray observatories.

CREDO user base

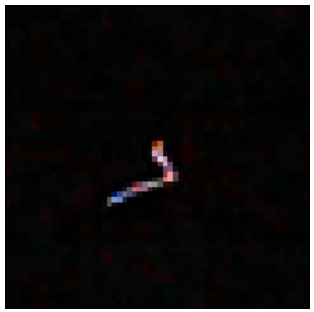
Statistics	
Users	$1.4 \cdot 10^4$
Candidate detections	$1.0 \cdot 10^7$
Devices	$1.5 \cdot 10^4$
Operation time (days)	$3.9 \cdot 10^5$ ($\sim$ 1050 years)

Introduction and Motivation

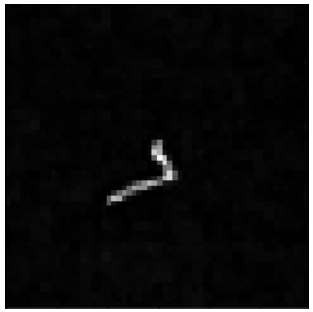
- The wealth of data collected by CREDO requires efficient means of rejecting the non-cosmic-ray noise and identification of signals attributable to extensive air showers.
- To address these problems we use:
  - Convolutional Neural Network-based trigger for artefact rejection and
  - Statistical Classifiers for signal morphological classification.
- Both approaches are based on supervised learning, so a representative subset of over 2300 CREDO images for training and validation was provided.

CNN based trigger

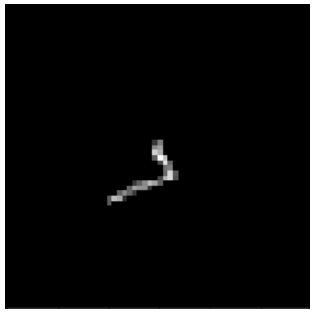
- Artefact rejection is preceded by three preprocessing steps:
  1. Grayscale conversion (cosmic rays have no color).
  2. Noise reduction by applying the adaptive thresholding.
  3. Wavelet transform (to amplify image features).
- Preprocessing steps are illustrated in worm image.



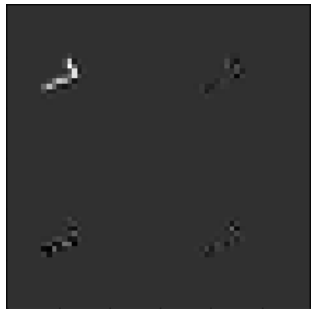
original



grayscale



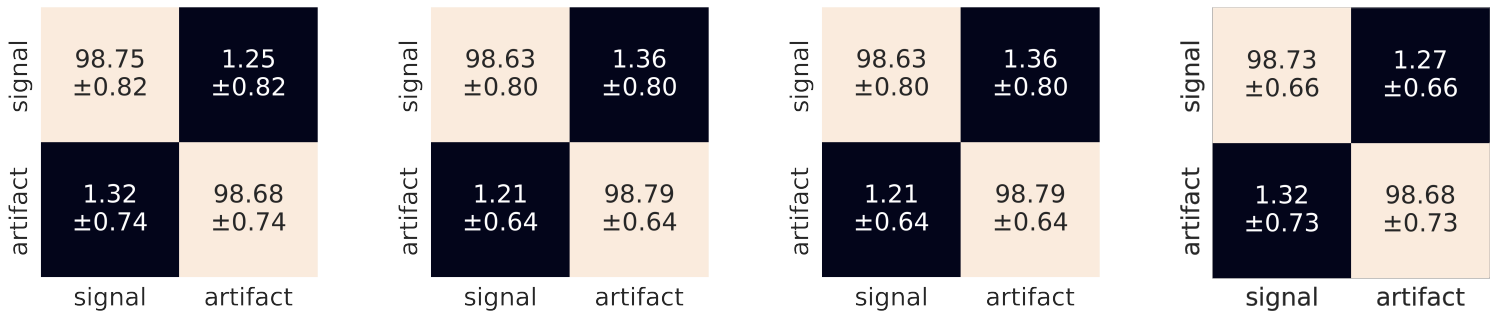
thresholding



wavelet

Trigger efficiency

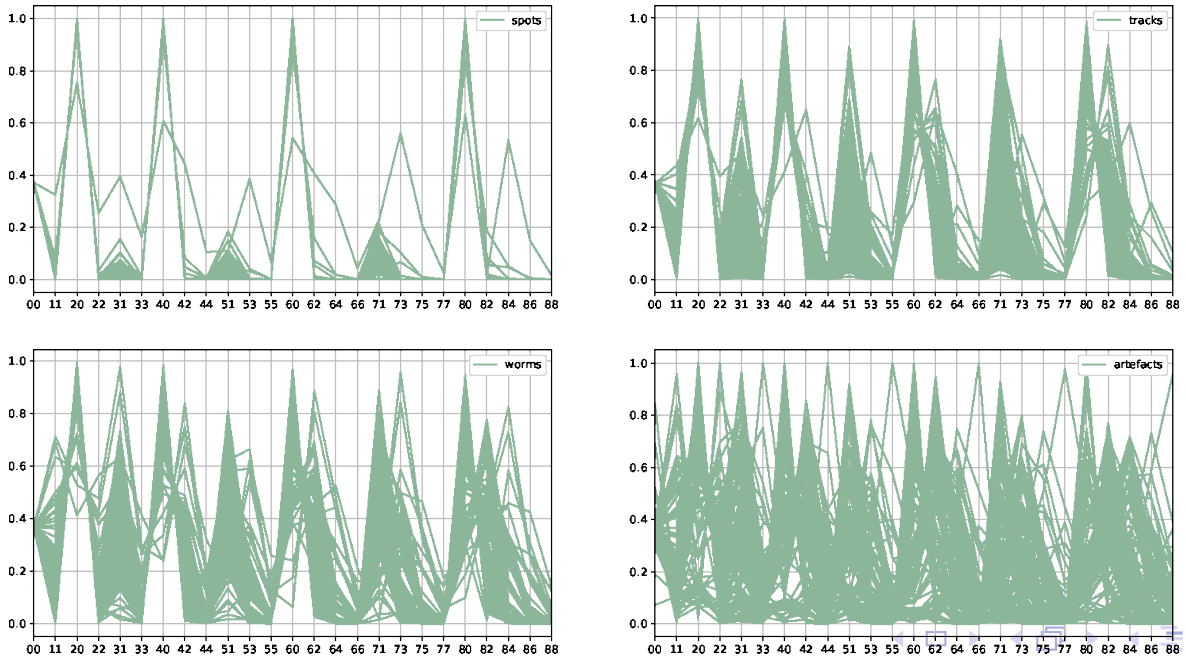
- Evaluation results for various wavelet input configurations (vertical - actual, horizontal - predicted):



- Conclusion: Trigger performance is not affected by the application of the wavelet transform.

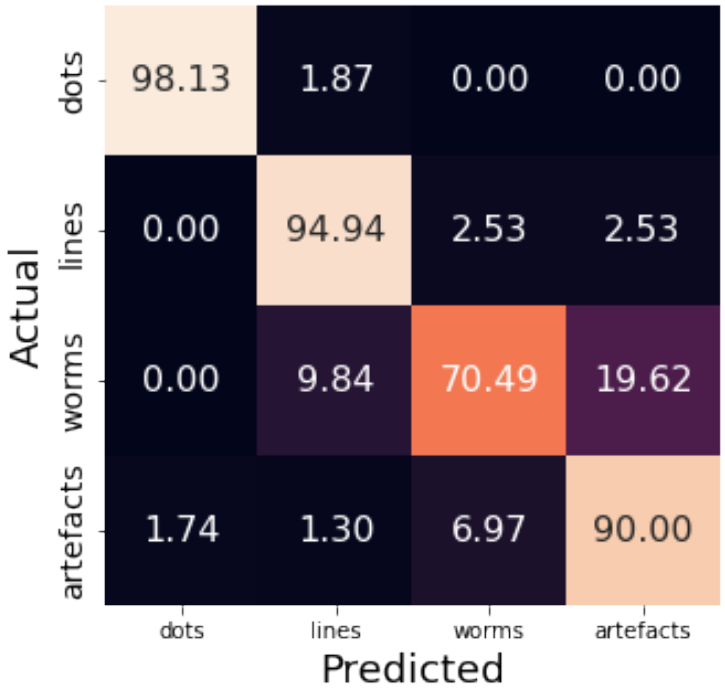
Zernike moments as feature carriers

- 4-classes are considered: spots, tracks, worms and artefacts.
- Zernike moments computed on thresholded input images are used as features.
- Difference between the classes can be observed on plots with Zernike moments.



4-class classification efficiency

- Among various considered classifiers the ν-SVC has exhibited the best performance.
- The general performance of ν-SVC is shown in figure below



Recognition rates

- Signal vs. Artefact: 99%, 99%,
- Dots vs. Tracks vs. Worms vs. Artefacts: 98%, 95%, 70%, 90%,

Further reading:

[1] Ł. Bibrzycki, D. Burakowski, P. Homola, M. Piekarczyk, M. Niedźwiecki, K. Rzecki et al., *Towards a global cosmic ray sensor network: CREDO detector as the first open-source mobile application enabling detection of penetrating radiation*, *Symmetry* **12** (2020) 1802.

[2] T. Hachaj, Ł. Bibrzycki and M. Piekarczyk, *Recognition of cosmic ray images obtained from cmos sensors used in mobile phones by approximation of uncertain class assignment with deep convolutional neural network*, *Sensors* **21** (2021) .

[3] M. Piekarczyk, O. Bar, Ł. Bibrzycki, M. Niedźwiecki, K. Rzecki, S. Stuglik et al., *CNN based classifier as an offline trigger for the CREDO experiment*, *Sensors (to be published)* (2021) .

[4] O. Bar, Ł. Bibrzycki, M. Niedźwiecki, M. Piekarczyk, K. Rzecki, S. Stuglik et al., *Zernike moment based classification of radiation signals in CMOS sensors*, *(in preparation)* (2021) .