Journal of Physics:
Conference Series

Convolutional neural network approach to event position reconstruction in DarkSide-50 experiment

To cite this article: A Grobov *et al* 2020 *J. Phys.: Conf. Ser.* **1690** 012013

Convolutional neural network approach to event position reconstruction in DarkSide-50 experiment

A Grobov and A Ilyasov for the DarkSide collaboration

National Research Center "Kurchatov Institute", Kurchatov sq. 1, Moscow, Russia
National Research Nuclear University MEPhI, Kashirskoe sh. 31, Moscow, Russia

E-mail: `ilyasovaid@yandex.ru`

Abstract. Neural networks are currently used in various fields of science and technology, as well as in experiments related to particle physics. DarkSide-50 is a two-phase (liquid and gas) argon TPC which has two main signals: scintillation in LAr (S_1 signal) and electroluminescence in GAr (S_2 signal) [1]. In the low-mass dark matter search only the more energetic second signal is used for position reconstruction [2]. As a result only events detected by seven central photomultiplier tubes are used for the analysis. Here we attempt to improve reconstruction using the convolutional neural networks (CNN).

1. Introduction

The DarkSide-50 is a two-phase liquid argon (LAr) time-projection chamber (TPC) which is located in Hall C of the Gran Sasso National Laboratory (LNGS) in Italy, at a depth of 3800 m.w.e. Its main purpose is to search for the rare nuclear recoils possibly induced by weakly interacting massive particles (WIMPs). The experiment is fully described in [1]. The DarkSide-50 detector has the following characteristic features, see figure 1:

(i) Argon target, corresponding to a 90% confidence level (C.L.) upper limit on the WIMP-nucleon spin-independent cross section of 6.1×10^{-44} cm^2 for a WIMP mass of 100 GeV/c^2 [3] and 90% C.L. exclusion limit above 1.8 GeV/c^2 for the spin-independent cross section of dark matter WIMPs on nucleons, extending the exclusion region for dark matter below previous limits in the range 1.86 GeV/c^2 [2].

(ii) (46.4 $\pm$ 0.7) kg. of LAr in the time projection chamber (TPC).

(iii) Gas layer between the LAr surface and the TPC anode.

(iv) Viewed by 38 photomultiplier tubes (PMTs), 19 each on the top and the bottom.

(v) It has wavelength shifter tetraphenyl butadiene (TPB), which absorbs 128 nm scintillation light and emits visible photons with peak wavelength 420 nm.

(vi) Voltage between cathode and grid creates vertical electric field in LAr to make ionization electrons drift up.

(vii) Pulse shape discrimination allows perfect β/γ background rejection;

(viii) To protect TPC from neutron and γ-ray events, TPC is enclosed in a liquid scintillator veto (LSV): 4.0 m-diameter stainless steel sphere, 30 t of borated liquid scintillator, 110 Hamamatsu R5912 8" PMTs [3].

(ix) LSV with TPC is immersed in water cherenkov detector (WCD) for muon shielding: 11 m-diameter, 10 m-high cylindrical tank, high purity water, 80 8” PMTs on the side and bottom of the water tank [3].

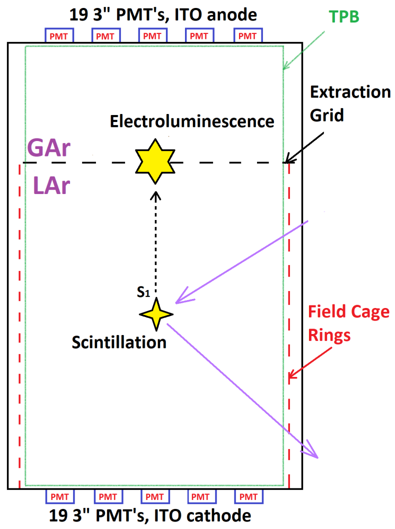

Figure 1. The DarkSide-50 experiment scheme: TPC scheme.

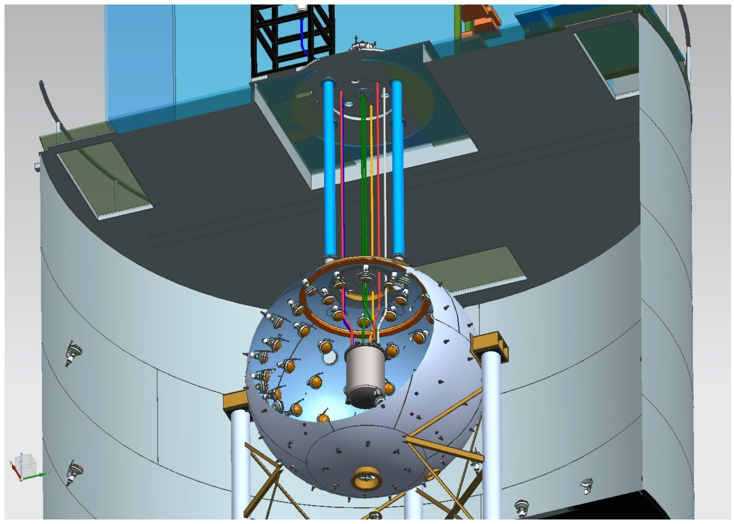

Figure 2. The DarkSide-50 experiment scheme: full detector system of DarkSide-50: WCD, LSV and TPC [1].

Position reconstruction algorithm is based on Monte Carlo (MC) simulation [3, 4]. Distribution of light in PMTs for every data event is used to calculate (x-y) position by minimizing χ^2 metric between it’s light response function (LRF) and all possible LRFs derived from MC events. The position resolution of this algorithm was estimated to be about 0.6 cm [3]. In case of low-mass dark matter [2] the recoil energy is too low to produce distinguishable scintillation signal, therefore S_2-only analysis is performed. Being unable to use S_1 signal the position of event is assigned to the center of the PMT, which detected maximum amount of light. Further fiducialization leaves us with volume under seven central PMTs [2].

2. Convolutional neural network technique

Convolutional neural networks is a deep learning algorithm that is often used for image recognitions and detection problems. These networks are based on *convolution* layers which

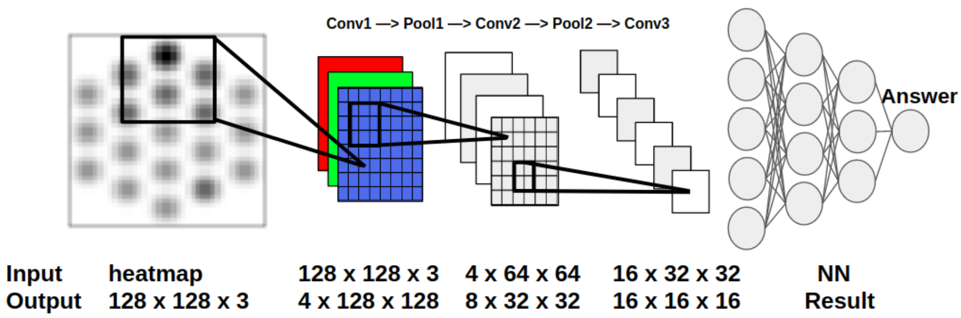

Figure 3. Convolutional neural network example for DarkSide-50 heatmap.

mathematically convolves the image in matrix representation with a set of kernel matrixes. Each kernel aims at underlining certain specific features of an input image. Output of the layer is then passed to the *pooling* layer, where dimensionality of the input matrix is further reduced. After the sequence of this layers the result is flattened and propagated through the fully connected layers.

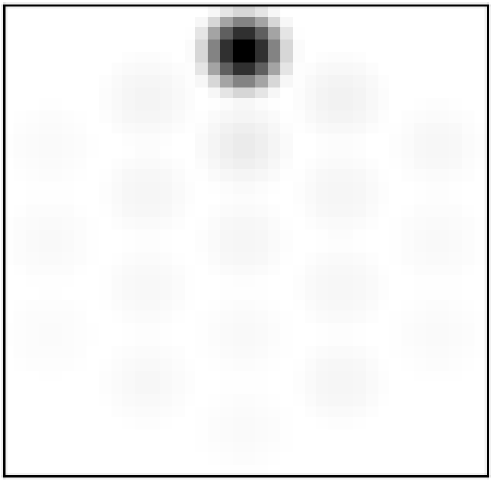

Figure 4. Photoelectrons distribution from S_2 signal in top 19 PMT's: original distribution.

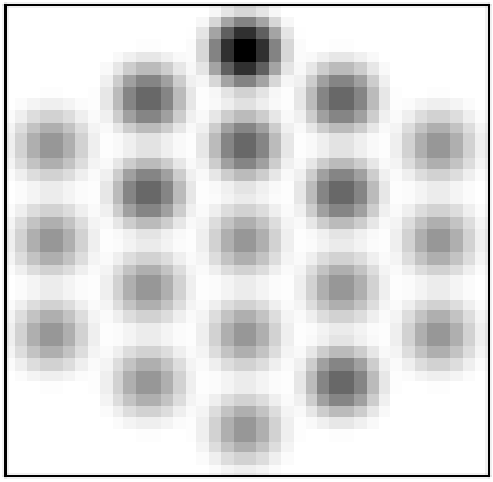

Figure 5. Photoelectrons distribution from S_2 signal in top 19 PMT's: scaled distribution.

The training process is as follows:

- Training samples were simulated with *G4DS* [5] package. G4DS is designed with a modular architecture in order to include a full description of all the detectors belonging to the DarkSide program [5]. We simulated $15 \cdot 10^3$ events in total. Every event is represented by image of photoelectron count in each PMT in form of heatmap.
- We split input data into training and testing samples and use only 33% of all events to train our model.
- Heatmaps are converted into matrix representation.
- Model is constructed as a plain stack of layers where each layer has exactly one input tensor and one output tensor [6].
- *Convolution layer* – performs convolution of input layer with different kernels to detect leading features of the input.
- *Max pooling layer* – downsamples the input representation by taking the maximum value over the defined window for each dimension along the features axis.
- output of convolution layers is collapsed into single dimension and passed to a fully connected layer.
- *Fully connected layer* – performs linear transformation and activation to predict X-Y coordinates. The weights of each neuron in the neural network are corrected after each training epoch.
- Optimization of parameters is done with stochastic gradient descent implemented from Adam algorithm [7] with mean squared error as a loss function.

Due to the fact that the S_2 signal is very strong and the PMTs are positioned very close to the LAr/GAr interface, a huge number of photoelectrons gets into only one photomultiplier tube. The distribution of light between neighbour PMTs is limited. This is a constraint for current reconstruction method, but we can avoid this logarithmically scaling the heatmaps.

3. Results

In this work we have the structure of CNN as on figure 3. PMT's heatmaps are transformed into a tensor of shape 128 $\times$ 128 $\times$ 3.

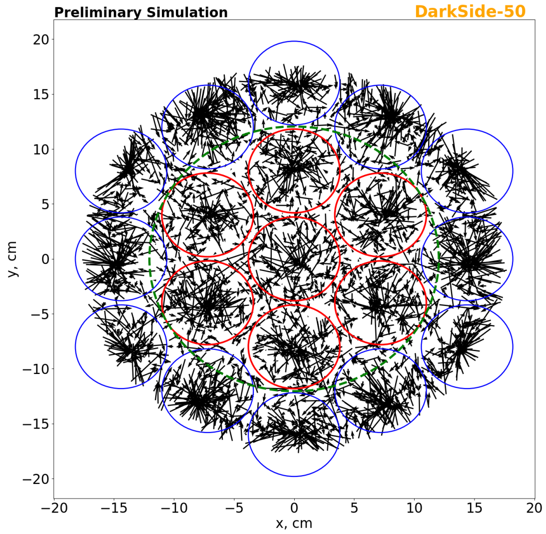

Figure 6. Plot, that shows reconstruction error: arrowtail and arrowhead correspond to the original and reconstructed position of the event: red circles correspond to 7 internal PMT's; blue circles correspont to 12 external PMT's; green circles corresponds to the volume of the detector in which the error was calculated.

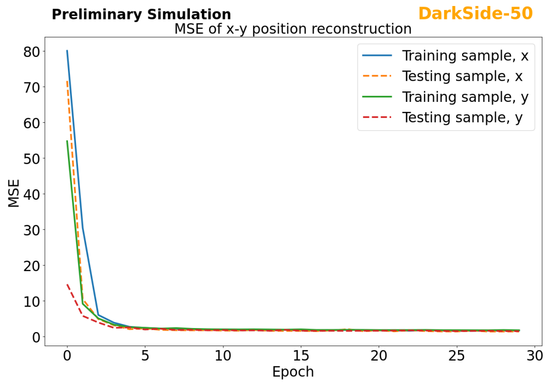

Figure 7. MSE value on each training and testing epoch.

Convolution layers result in an output tensor of 16 $\times$ 16 $\times$ 16 size. The result is propagated through a 4 fully connected layers (4096 $\times$ 512 $\times$ 128 $\times$ 32) and finally to the output.

X and Y coordinates are reconstructed separately. Figure 7 shows the mean square error for training and testing samples at each epoch. We see that error is stable during the last $\approx$ 30 epochs.

Model performance is shown in figure 6. All events tend to reconstruct towards PMT centers. We conclude that there is room for improvement, but even simple CNN model is able to reconstruct position fairly well in the fiducial volume. Total MSE on train sample $\approx$ 1.5 cm, total validation MSE $\approx$ 2 cm for both X and Y reconstruction. Mean error on test sample for all events $\approx$ 0.33 (0.03) cm and for 7 central PMT's $\approx$ 0.4 (0.06) cm for reconstruction X (Y) coordinates respectively. The reason why there is a difference in coordinates reconstruction is left for further investigation. The result we have achieved can be used to repeat low-mass analysis and update existing limits.

Acknowledgments
The work was supported by Russian Science Foundation Grant No 16.12.10369.

References
[1] Agnes P *et al.* 2015 *Phys. Lett.* B **743** 456–466
[2] Agnes P *et al.* 2018 *Phys. Rev. Lett.* **121** 081307
[3] Agnes P *et al.* 2018 *Phys. Rev.* D **98** 102006
[4] Brodsky J P 2015 xy position reconstruction in darkside-50 Tech. rep. Fermi National Accelerator Lab. (FNAL), Batavia, IL (United States)
[5] Agnes P *et al.* 2017 *J. Instrum* **12** P10015
[6] Chollet F *et al.* 2015 Keras URL `https://keras.io`
[7] Kingma D P and Ba J 2014 Adam: A method for stochastic optimization *3rd International Conference on Learning Representations,* ICLR 2015 ed Bengio Y and LeCun Y (*Preprint* `1412.6980`)