

# BESIII Drift Chamber Tracking with Machine Learning

Yao ZHANG<sup>1,\*</sup>, Ye YUAN<sup>1</sup> and Qiumei MA<sup>1</sup>

<sup>1</sup>Institute of High Energy Physic, Beijing Yuquan Road 19B, China

**Abstract.** The tracking efficiency and the quality for the drift chamber of the BESIII experiment is essential to the physics analysis. The tracking efficiency of the drift chamber of BESIII is high for the high momentum tracks but still have room to improve for the low momentum tracks, especially for the tracks with multiple turn. A novel way to use a convolutional network called U-Net network is represented to solve the identification of the first turn's hits for the multiple-turn tracks.

## 1 Introduction

### 1.1 The BESIII experiment

The Beijing Spectrometer III (BESIII[1], Fig.1) has been running at the Beijing Electron Positron Collider II (BEPCII) for Tau-Charm physics since 2008. The tracking detector of the BESIII is a Multilayer Drift Chamber (MDC)[2]. The tracking efficiency for the high transverse momentum is high but lower when the  $\cos\theta$ (where  $\theta$  is the dip angle of the track ) is small for the low transverse momentum tracks (Fig. 2).

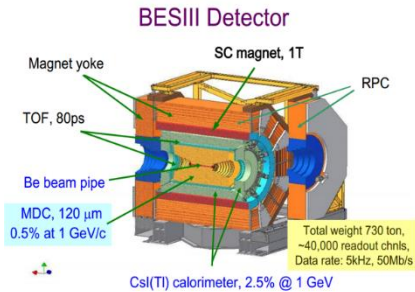


Fig. 1. The BESIII Detector.

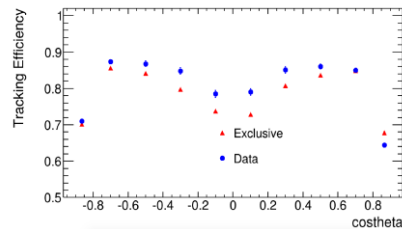
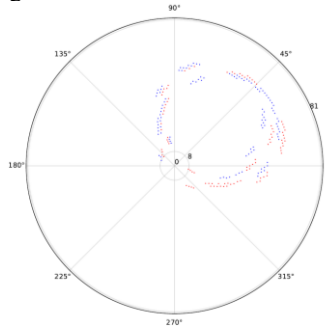


Fig. 2. The tracking efficiency of the  $\pi^-$  along  $\cos\theta$

\* Corresponding author: [zhangyao@ihep.ac.cn](mailto:zhangyao@ihep.ac.cn)

## 1.2 Multi-turn track finding of the drift chamber

Some of the curling tracks with multiple turn which  $|\cos\theta| < 0.2$  have bad tracking efficiency and quality as Fig. 2 shows. For the multi-turn tracks, the hits from different turns will be adjacent to each other and even overlap as Fig. 3 shows. The tracking efficiency is low due to the effect of the non-first turn's hits during tracking procedure. The solution to remove the non-first turn hits in traditional tracking algorithm is to apply a cut on the distance between the track and the MDC wires. The disadvantage of the traditional method is that the track should be reconstructed with the contamination of the non-first turn hits and worsen the track parameter which is used for the selection of the hits. In our work, the first turn hits can be identified without the knowledge of the track.



**Fig. 3.** An event display of the multi-turn track in the MDC in polar coordinates. Dots are the wire position at the endcap with hit. Red dots are the hits from the first turn and blue dots are the hits from the non-first turn.

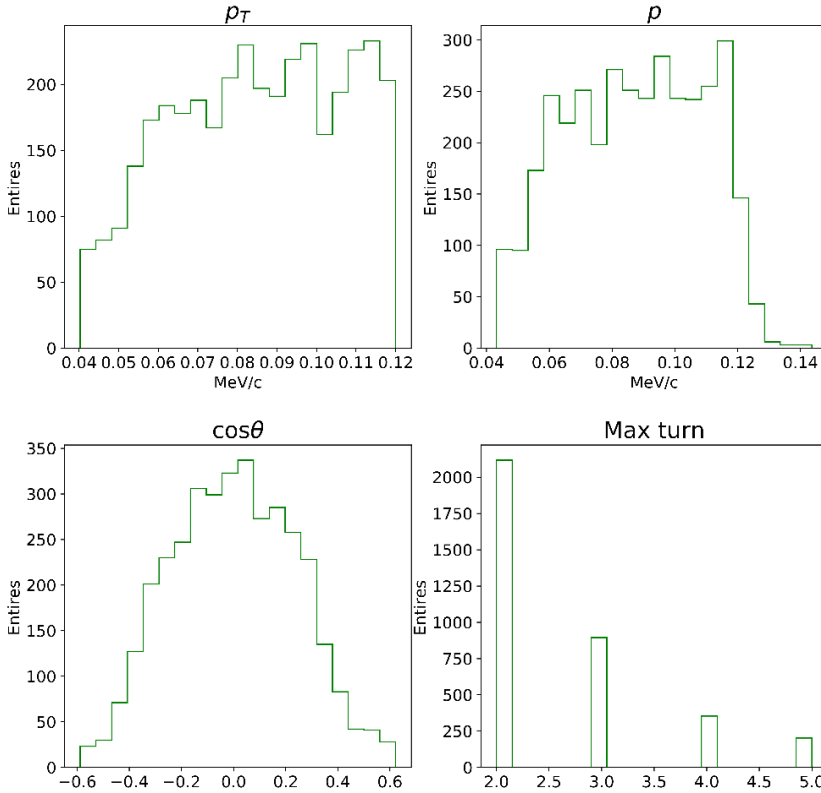
In our work, we use the technology of the supervised semantic segmentation of the deep learning to do the hit level classification for the single track. The aim of this work is to separate the first turn's hit from the non-first's turn hits for the multi-turn tracks.

## 2 Multi-turn hits separation with U-Net

The track finding of the tracking detector is a problem of pattern recognition which is very suitable for the utilization of the machine learning or deep learning technique. In the work of this paper, the hits from one event in drift chamber are mapped as the elements in a 2-dimensional matrix which will be described in detail in the session 2.1. The problem of identifying the hits between the first turn and non-first turn became a problem of training a convolutional neural network to do the pixelwise semantic segmentation for the matrix of each event.

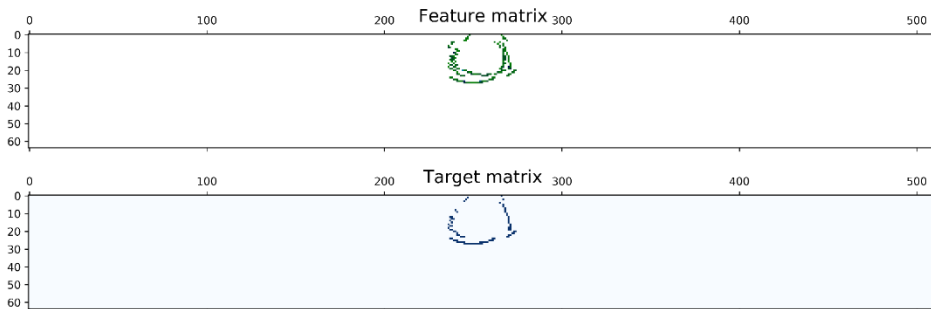
### 2.1 The training sample

To learn the behavior of the multi-turn tracks, the training sample is the multi-turn curling tracks generated with Monte-Carlo which are at small dip angle and various transverse momentum. The particle type is single  $\mu^-$  and single  $\mu^+$ . The tracks with transverse momentum less than  $120 MeV/c$  which will curl in the MDC are generated. Only the tracks of more than one turn have been selected for the training, validation and testing. The momentum, angle and turn number distribution of the training set show in the Fig. 4.



**Fig. 4.** Distributions of the training dataset

The MDC has 43 layers and the cell number of each layers range from 40 to 288. To make the shape fit between input and output, the rows and columns of the matrix are set to be the powers of 2. The cells of the drift chamber for each layers were top centered mapped into the corresponding row of a 64 x 256 matrix. The feature set to be the raw time of each hit after correction with the event start time. The feature of the cell without hits were set with a default value of 2000ns. The truth turn number from Monte-Carlo simulation used as target. The hits from the first turn were marked as one. The hits from non-first turns and cells without hit were set to be the default value of zero. And zero also treated to be the default value of the target. Dummy elements in the matrix were set to be the default values.

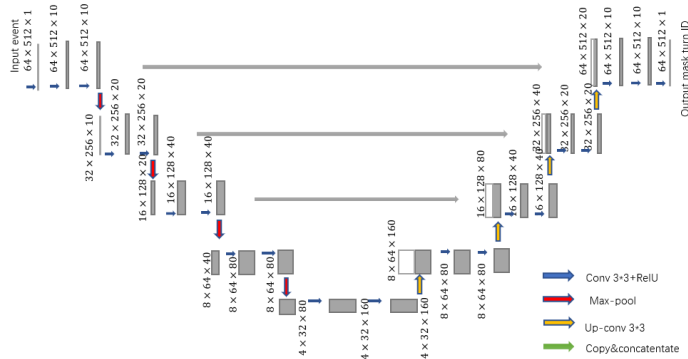


**Fig. 5.** The feature and target matrixes of one track

## 2.2 Network setup and training

### 2.1.1 Network model

U-Net[3] is a convolutional network for the pixel level semantic segmentation which is widely used in the computer vision. Following is the U-Net model used in this work.



**Fig.6.** The feature and target matrixes of one track

### 2.1.2 Loss function

The loss function used in this work is a Binary Cross Entropy (BCE) plus a dice coefficient loss. Dice coefficient loss as a deep learning loss function for the highly unbalanced segmentations. It was chosen because of the unbalanced number of the elements between first turn and non-first turn.

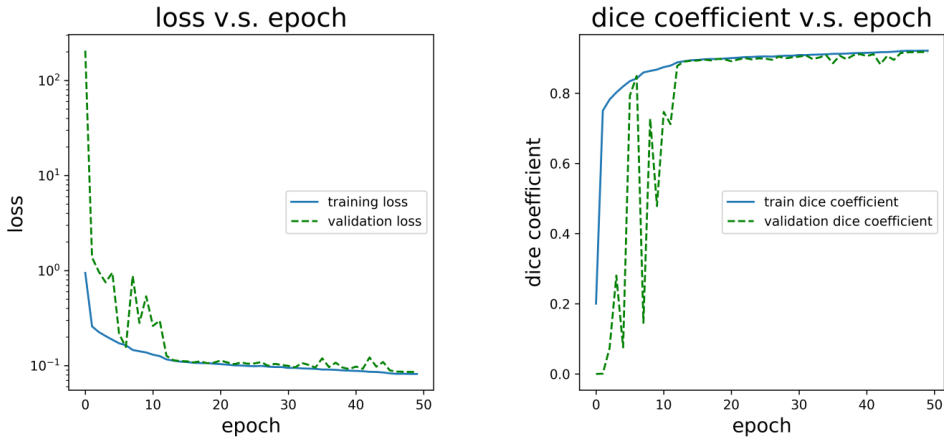
$$Dice\ coefficient = \frac{2|X \cap Y|}{|X| + |Y|} \quad (1)$$

$$loss = 1 - Dice\ coefficient + BCE \quad (2)$$

Where, X is the target matrix and Y is the prediction matrix. In this work, |X| and |Y| are approximately equal with the number of the true first turn's hits and the predicted first turn's hits respectively, |X ∩ Y| is approximately equal with the number of the hits marked as the first turn hits and predicted as the first turn hits.

### 2.1.3 Training

A sample of 10 thousand tracks have been used for the training. The dataset is separated into 7500 tracks for the training and 2500 tracks for the validation. The batch size is 64 and the epoch number is 50. Following are the training loss and the model accuracy along epoch number. The training has converged after about 15 epochs for the validation dataset.



**Fig.6.** Loss and dice coefficient v.s. epoch

### 3 Performance and conclusion

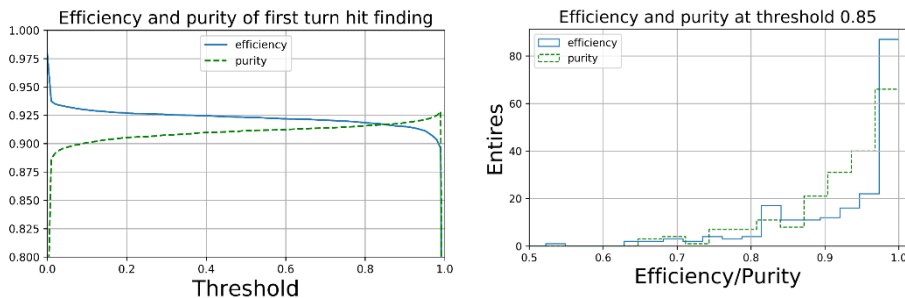
The efficiency and purity have been evaluated for the single  $\mu^-$  track with validation sample. The definition of the efficiency and the purity show as following:

$$Efficiency = \frac{TP}{P} \tag{3}$$

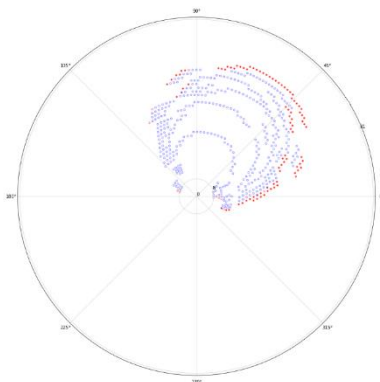
$$Purity = \frac{TP}{P'} \tag{4}$$

Where TP is the number of the predicted first turn and marked as first turn by Monte-Carlo truth, P is the number of the true first turn hits and  $P'$  is the number of the hits predicted as first turn.

The ROC curve of the average hit efficiency and purity in Fig. 7 shows that at the threshold of about 0.85 the prediction gives the efficiency and purity are about 91%. And more than half of the events have efficiency and purity greater than 95%. An event display in Fig. 8 shows most of the hits from first turn have been predicted correctly.



**Fig.7.** Left is the ROC curve of the efficiency and purity. Right is the efficiency and purity on threshold at 0.85



**Fig.8.** An event display of a multi-turn track after prediction by U-net network in a polar coordinate. Red dots are the first turn hits predicted as first turn. Red circles are the first turn hits predicted as non-first turn. Red dots with blue circles are the non-first turn hits mis-identified as first turn.

A deep learning effort has been performed to solve the problem of the multi-turn curling track finding for the BESIII drift chamber. The preliminary performance is satisfied and promising. When threshold is at 0.85, the average efficiency and purity for the first turn's hits is about 91%. We intend to investigate the separation of the hits for each track object using instance segmentation in deep learning in the future.

## References

1. BESIII Collaboration, Nuclear Instruments and Methods in Physics Research Section A, **614**(3) (2010)
2. W. L. Yuan, et al., Chin. Phys. C, **40** (2) 026201 (2016)
3. Olaf Ronneberger, Philipp Fischer, Thomas Brox, arXiv: 1505.04597 18 May (2015)