Class Separation and Parameter Estimation with Neural Networks for the XEUS Wide Field Imager

J. Zimmermann[a], C. Kiesling[b] and P. Holl[c].

[a]Forschungszentrum Jülich GmbH, Zentrallabor für Elektronik, 52425 Jülich and
Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München

[b]Max-Planck-Institut für Physik, Föhringer Ring 6, 80805 München

[c]PNSensor GmbH, Römerstraße 28, 80803 München

The X-ray space telescope XEUS is the proposed follow-up project to ESA's cornerstone mission XMM-Newton which is now in orbit. To face the high data rate from the pixel detector and to improve event processing neural networks are under study to be integrated into the electronics on board (online) and to serve as analysis tool on ground (offline). For two applications results are presented. First as a typical online application, the separation of single photon events from pileup: here the unwanted event topologies are separated from useful ones belonging to a single X-ray photon. Second a typical off-line application, the estimation of the incident position of a photon: here the charge splitting (i.e. signal charges are collected by two or more adjacent pixels) can be used to determine a precise incident position of a photon. The neural network results are compared with standard methods.

1. Introduction

The X-ray Evolving Universe Spectroscopy (XEUS) mission [1] is a potential follow-up project to ESA's cornerstone X-ray Multi-Mirror mission (XMM-Newton).

The science goal of the XEUS mission is the hot universe at far redshift, the first massive black holes, the first galaxy groups and the evolution of heavy elements. The satellite will probably be launched after 2012.

The collection area will be 6 m^2 (XEUS I) and 30 m^2 (XEUS II), corresponding to an increase in photon count rate of factor 40, respectively 200, compared to XMM-Newton. For high throughput imaging X-ray spectrometry, an active pixel sensor is being developed [2,3]. It will act as Wide Field Imager (WFI) and match the optical properties of the X-ray mirror optics due to its pixel geometry, energy response, low noise performance, and fast readout.

To face the resulting data rate, neural networks are under study to be integrated into the electronics on board for efficient background suppression. The necessary speed of the neural calculation is achieved by a hardware implementation (in FPGA technology) which makes use of the inherent parallelism of feed-forward neural networks[1].

Neural networks can also be used to improve event processing by estimating properties of photon events like the incident position. That may also be done on board or in the offline analysis on ground.

The data which were used to train and evaluate the neural networks were generated with a Monte Carlo simulation code.

2. Neural Networks Framework

We developed a general framework for training, application and evaluation of feed-forward neural networks in the ROOT-environment [4]. The neural network code is based on classes from J.P. Ernenwein [5].

The implemented feed-forward networks are trained using the backpropagation algorithm.

[1]In each layer all the neurons can be processed in parallel which results in a two step process for a so-called three layer neural network.

This type of neural networks has proven to be the appropriate neural architecture for class separation and parameter estimation in many applications [6]. We created an "automatic" training procedure where the two parameters of the algorithm (learning rate and momentum parameter) are changed dynamically during the training to make it fast and complete. Figure 1 shows the evolution of the cost function under the same starting conditions for three different training methods.

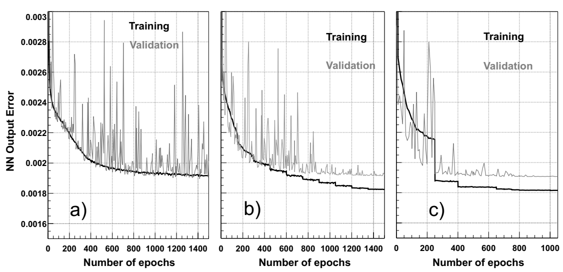

Figure 1. a) Fixed number of epochs, constant learning parameters, b) Fixed number of epochs, decreasing learning parameters, c) "automatic" training procedure

3. Single Photon Recognition

For the astrophysical data analysis mainly the energy and incident position of X-ray photons are of interest. These values can only be determined from an event[2] if it is isolated from other events (the contrary is called a pileup). To examine the capabilities of neural networks to reject events which do not represent a single photon we trained networks to separate single photon events from pileup events. Examples for these two classes are shown in figure 2.

The input for the neural network consists of four values representing the split charges. Since a single photon event cannot occupy more than a 2×2 grid in the XMM/XEUS pixel-geometry,

[2]An X-ray photon usually illuminates a group of up to four neighbouring pixels, an event is called split-event if the charge is distributed over more than one pixel.

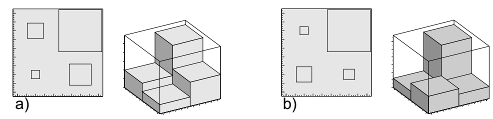

Figure 2. a) Single photon vs. b) Pileup event

only the pileups within that size need to be processed.

The standard algorithm currently used for this problem in the XMM-data analysis [7] uses four patterns which are valid for single photon events, as shown in figure 3. Additionally, a total-charge-cut is applied[3].

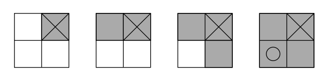

Figure 3. Patterns for the standard approach of single photon recognition: Grey pixels are illuminated, a cross marks the position of the maximal charge, a circle the position of the minimal charge. The rotated patterns are also valid as single photons.

The results of both methods are shown as efficiency versus rejection curves in figure 4. While the cut value in the network output can be varied, the parameters for the standard algorithm are the threshold value[4] and the total-energy-cut. As can be seen from figure 4, we observe a much better capability of background suppression for the neural method.

4. Position Estimation

An optimal angular resolution of the detector is important for the astrophysical application. We therefore want to determine the incident position of the photon with subpixel precision. For this we center the maximal charged pixel of a pho-

[3]According to the maximum energy a single photon could have, a total charge exceeding this limit is considered as a sign for pileup.

[4]The threshold value is applied to recognize a pixel as illuminated.

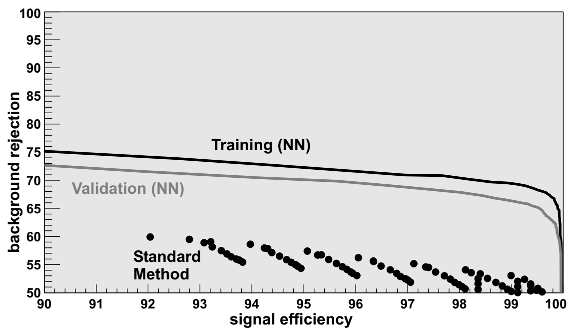

Figure 4. Efficiency vs. rejection for standard and neural network solution

ton event in a 3×3 grid and try to derive a position information from the distribution of the split charges. An example is shown in figure 5.

These 3×3 grids serve as input for the neural network while the output should be the distance of the incident position to the left border of the central pixel as indicated in figure 5.

We compare the neural network results with the standard approach where the center of mass (COM) of the split charges is calculated. Afterwards a correction table is applied to these values, which has been calculated as the difference between the COM result and the true position for simulated events. By this it takes into account the Gaussian shape of the charge clouds.

In the results which are shown as error distributions $x_{est} - x_{true}$ in figure 6 one can observe a 15% smaller error for the neural network than for the

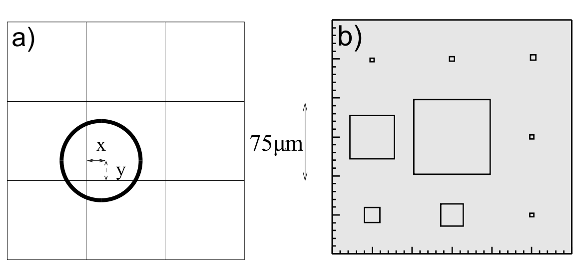

Figure 5. Deriving the position from split charges: a) Position of the center of the charge cloud, b) Resulting charge distribution

corrected center of mass method (CCOM). This makes the neural network method also attractive for offline analysis.

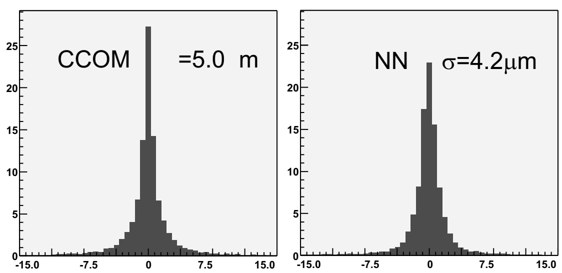

Figure 6. Error distributions for standard and neural network solution

REFERENCES

1. European Space Agency, Research and Scientific Support Department, "The X-ray Evolving Universe Spectroscopy Mission", `http://astro.esa.int/SA-general/Projects/XEUS/main/xeus_main.html`.
2. P. Holl et al., "Active Pixel Matrix for X-Ray Satellite Missions", IEEE Transactions on Nuclear Science, Vo. 47, No.4:1421-1425 (2000).
3. P. Holl et al., "Active Pixel Sensors for Imaging X-Ray Spectrometers", proceedings of SPIE's Astronomical Telescopes and Instrumentation, Hawaii 2002, to be published.
4. R. Brun, F. Rademakers et al., "ROOT, An Object-Oriented Data Analysis Framework", `http://root.cern.ch`.
5. J.P. Ernenwein, "Neural Network NeuNet", `http://e.home.cern.ch/e/ernen/www/NN/index.html`.
6. C. Kiesling et al., "The H1 Neural Network Trigger Project", ACAT 2000, VII International Workshop, Batavia, Illinois, AIP Conference Proceedings Vo. 583, p. 36-44.
7. K. Dennerl, U.G. Briel, F. Haberl, N. Krause, M. Popp, V.E. Zavlin, "Determination and correction of the charge transfer efficiency of the pn-CCD camera", SPIE's 44-th Annual Meeting and Exhibition, 3765-24 (1999).