Neural network based cluster reconstruction in the ATLAS pixel detector

K. Prokofiev\textsuperscript{a}, K. E. Selbach\textsuperscript{b}

for the ATLAS Collaboration
\textsuperscript{a}New York University, USA, \textsuperscript{b}University of Edinburgh, Edinburgh, United Kingdom
E-mail: kselbach@cern.ch

Abstract. The ATLAS Pixel detector is currently measuring particle positions at 8 TeV proton-proton collisions at the LHC. In the dense environment of jets with high transverse momenta produced in these events the separation between particles becomes small, such that their respective charge deposited are reconstructed as single clusters. A Neural Network (NN)-based clustering algorithm has been developed to identify such merged clusters. By using all cluster information, the NN is ideal to estimate the particle multiplicity and for each of the estimated number of particles, the position with its uncertainty. As a result of the NN reconstruction, the number of hits shared by several tracks is strongly reduced. Furthermore, the impact parameter improves by about 15\% which indicates boosted prospects for physics analysis.

1. Introduction
The ATLAS silicon pixel detector determines the position of passing charged particles. It consists of three layers where each layer consists of a barrel and two endcap, one at each end of the barrel. Mostly 50\,\mu m \times 400\,\mu m pixels are used in the barrels which are at a radius of 50.5\,mm, 88.5\,mm and 122.5\,mm [1]. In the standard clustering algorithm, pixels with deposited charge are grouped into clusters if they have a common edge or a common corner. By using the charge deposited in each pixel, measured from the time over threshold (ToT) by the sensors, the resolution of the particle’s position can be improved using charge interpolation. The standard reconstruction algorithm uses a linear interpolation between the first and last set of pixels in a cluster. In a densely populated track environment (e.g. a jet) two or more particles might deposit charge in adjacent pixels whereby shared clusters are produced. As shown in Figure 1, it is likely that tracks of jets with $p_T$\,>\,1 TeV produce merged clusters. Thus, the aim of the Neural network (NN) application to the cluster reconstruction algorithm is to determine more precise cluster positions and improve the two-particle separation. This is achieved by using the detailed information of the cluster shape and identifying merged clusters produced by nearby particles.

2. Neural Network algorithm
A NN has the ability to solve pattern recognition problems and handle non-linear correlations between input variables. Furthermore, these inputs are differently weighted in the hidden layers of the NN to finally determine the output. These properties of the NN meet the requirements of
the cluster reconstruction which has to deal with many cluster properties which are alone nearly meaningless (e.g. the charge of a single pixels). When these cluster properties are combined and put in context (e.g. knowing the charges of the adjacent pixels), these variables contain all information required for successful pattern recognition. The following data are used as inputs for the NN:

- Matrix with $7 \times 7$ pixels containing the deposited charges within the original cluster. The cluster is centered in the matrix by using the charge weights determined by ToT;
- Vector containing information of the longitudinal size of the pixel. It distinguishes between standard length of 400 $\mu$m and long pixels of 600 $\mu$m;
- Estimated direction of the traversing charged particle.

As a results the NN algorithm provides an estimate of the number of charged particles passing through each cluster. Furthermore, for each of this estimated number of particles, the position with its uncertainty is determined in two dimensions. A selection of the inputs and outputs of the NN algorithm are shown in Figure 2. In particular, a $6 \times 4$ extract of the matrix is displayed containing the charges deposited in each pixel represented by the color code. In the lower left corner, it is highlighted that the probability of two particles passing through this selected cluster is higher than any other possible particle multiplicity. Therefore, two estimated particle positions are given by the blue markers. By comparing them to the positions (pink marker) and the directions (red lines) of the true particles, the functionality of the NN algorithm is depicted.

Different sets of NN are used to determine the particle multiplicity, the particle position and the errors on the particle position. In fact, all of the NN use the same three inputs specified above. To train the NN, mixed samples of $t\bar{t}$ and high $p_T$ dijet events ($140 < \text{jet } p_T < 560 \text{ GeV}$) are used.
3. Results

As shared clusters are produced in densely populated track environment, some clusters are associated to multiple tracks. These shared hits on-track are less likely when using the NN algorithm as Figure 3 illustrates. This is shown for the innermost Pixel layer where the NN algorithm can produce the greatest improvement with respect to the other layers. As the innermost layer is as close as 50.5 mm to the interaction point the separation of tracks is especially small. The resulting large amount of shared clusters offers potential to the NN to develop its full abilities. Apart from this, the NN reconstruction performs similarly well in the other layers. For tracks with high $p_T$, the intrinsic detector resolution is critical as the impact of multiple scattering is reduced. Especially in this environment the NN can provide a crucial improvement. To display this, tracks with $p_T > 100$ GeV are selected for the Figure 3.

Figure 4 illustrates the cluster residual, which is defined as the difference between the cluster position and the extrapolated track position. The shape of the cluster residual is a double-peak structure for the standard clustering of a cluster with a size of 4 pixels. Clusters of this size are mainly produced by delta rays. Due to the non-linear treatment, the NN algorithm especially improves the particle position resolution in cases where delta rays are emitted by the traversing particle. Mainly due to this optimized treatment of delta rays, the double-peak vanishes when applying the NN clustering. In addition, the width of the cluster residual distribution decreases significantly which indicates an improved resolution. This improvement can be also observed for other clusters sizes or for the residual in $Z$. The increased cluster position resolution has direct consequences on the impact parameter. As explained earlier in the chapter, the NN algorithm most strongly improves the cluster position in the innermost layer. The innermost cluster position determines most considerably the impact parameter resolution. So, as shown in Figure 5, the longitudinal impact parameter is improved by $\sim 15\%$ for tracks with high $p_T$ for which the improvements by the NN reconstruction are particularly striking. Also, the transverse impact parameter shows a similar improvement with the NN algorithm. The improvement of the cluster position resolution has, therefore, a direct impact on primary and secondary vertex reconstruction especially relevant for heavy flavor tagging.

Figure 3: The number of shared hits on-track of the innermost Pixel layer reconstructed with and without the NN-based clustering for tracks with $p_T > 100$ GeV of $t\bar{t}$ samples

Figure 4: The cluster residual in the $r\phi$-direction for 4-pixel clusters reconstructed with the standard algorithm and the NN clustering algorithm for $t\bar{t}$
5: Simulated impact parameter resolution with and without NN-based clustering for tracks with $p_T > 100$ GeV of a $t\bar{t}$ sample

4. Conclusion

The NN algorithm significantly improves tracking performance for physics analysis. The improvements coming from the NN reconstruction will appear even more relevant when reaching the design luminosity of the LHC and in future upgrades, because of the higher particle density and the greater emphasis on high $p_T$ jets. During the 2013 shutdown, the ATLAS pixel detector will be upgraded with the insertion of a new layer closer to the interaction point ($R = 33.25$ mm) [2]. In this extremely dense environment the improved two particle separation power of the NN algorithm will boost the detector capability extending the physics potentiality of the ATLAS experiment.

5. References

[1] The ATLAS Collaboration, "The ATLAS experiment at the CERN Large Hadron Collider", *Journal of Instrumentation*, 3(08):S08003, 2008.

[2] M. Capeans et al., "ATLAS insertable b-layer technical design report", CERN-LHCC-2010-013, ATLAS-TDR-019, 2010.