

Towards the Increase in Granularity for the Main Hadronic ATLAS Calorimeter: Exploiting Deep Learning Methods

Philipp Gaspar, on behalf of the TileCal Collaboration

COPPE/Signal Processing Lab - Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil

E-mail: philipp.gaspar@cern.ch

Abstract. During the second phase upgrade program developed for LHC and its experiments, the main hadronic calorimeter of ATLAS (TileCal) will replace completely its readout electronics, but the optical signal pathway and detector will be kept unchanged. During the R&D studies for the upgrade, initial analyses for improving the calorimeter granularity were made. A granularity improvement could be achieved through the introduction of Multi-Anode Photomultiplier Tubes (MA-PMTs) into the calorimeter readout chain, together with applications of image processing algorithms for identifying sub-regions on calorimeter cells. This paper presents the latest results from using a Generative Adversarial Network (GAN) to generate synthetic images, which simulate real images formed in the MA-PMT. After the classification of cell sub-regions, preliminary results show a classification accuracy of more than 98% on the experimental test set.

1. Introduction

The Large Hadron Collider (LHC) accelerates and collides proton beams at an event rate of 40 MHz, with up to 13 TeV of center-of-mass energy. ATLAS [1] is the largest LHC detector, and it was designed as a general-purpose particle detector to study the extensive LHC physics program; going from the search of the Higgs boson, to the search for extra dimensions and dark matter particles.

LHC and its experiments have been developing an extensive upgrade program, which is split into two phases: Phase-I, going from 2019 to 2020, and Phase-II that will occur between 2024 and 2026. During the Phase-II upgrade, the main hadronic calorimeter of ATLAS (TileCal) [2] will redesign its front-end and back-end electronics, leaving the detector and optics unchanged. During the initial R&D studies for the TileCal Phase-II upgrade, a study was developed to improve the detector granularity by modifying only the calorimeter readout, without changing its mechanical structure. Particle jets with high transverse momentum tend to deposit their energies in the outer layers of TileCal. Therefore, dividing the actual calorimeter cells into new sub regions would improve the reconstruction of momentum, mass, transverse energy and angular position of those jets, allowing future analysis benefits from a finer-grained granularity detector.

This paper introduces a deep learning method to address the granularity improvement in TileCal. The method combines the use of generative adversarial models to create new synthetic

data in order to leverage the statistics, together with a classification algorithm used to assign each event to a particular topological region of the detector. We propose to use a variant of a Generative Adversarial Network (GAN) [3], in order to simulate the interaction of particles within a calorimeter cell, and, thereafter, leveraging the amount of statistics for the final classification model development. A variety of GANs has been used in the physics community to create fast experiment simulation in several analyses, like simulation for particle showers [4] and convergence maps in cosmology [5].

In this paper a variant of a GAN model based on a deep neural network (DCGAN) [6] is used to obtain a substantial increase in the statistics of the physical process being studied. Those synthetic events feed a supervised deep learning approach based on Convolutional Neural Networks (ConvNet) [7], which will be used for mapping the signal image information onto two regions from a particular calorimeter cell.

This paper is organized as follows. Section 2 describes the ATLAS calorimetry system, while Section 3 addresses possible improvement in the TileCal granularity even keeping unchanged its mechanical structure. The preliminary results of the proposed algorithm are presented in Section 4, and conclusions are derived in Section 5.

2. The Tile Calorimeter

The ATLAS calorimeter system is split into two detectors, the Liquid Argon (LAr) Calorimeter, which is responsible for measuring the energy of electromagnetic particles, and the Tile Calorimeter (TileCal), which measures the energy of hadronic particles. Both calorimeters are segmented in depth into layers, with fixed granularity (number of cells per layer). A schematic view of the ATLAS calorimeters together with a TileCal sketch is shown in Figure 1.

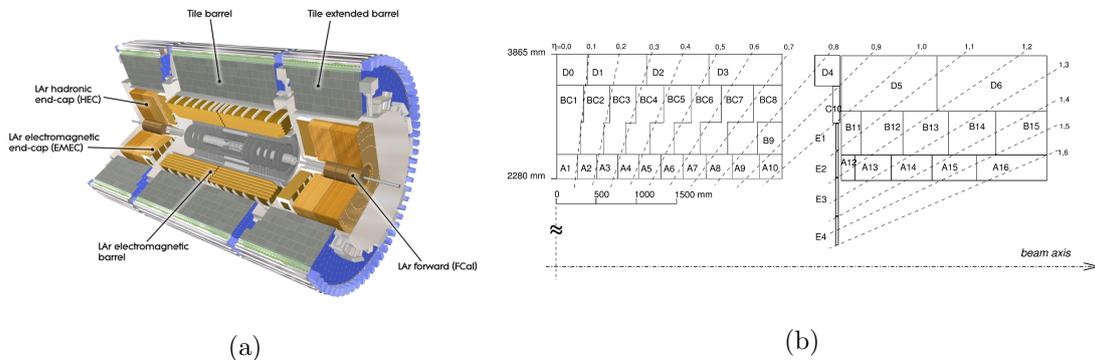


Figure 1: ATLAS calorimeters (left). TileCal cell layout (right).

TileCal consists of one Central Barrel (LB) and two Extended Barrels (EB) that together cover a pseudo-rapidity region of $|\eta| < 1.7$ with each cylinder segmented into 64 modules. TileCal is a sampling calorimeter and each cell is made of alternating layers of iron plates and scintillating tiles. The scintillating tiles are readout by wave-length shift fibers, which deliver the light to a Photomultiplier Tube (PMT) placed in the outer radius of the iron structure. Each PMT reads the signal from several tiles that are grouped into cells of different sizes, depending on their radius and pseudo-rapidity. The calorimeter modules are segmented into three longitudinal layers (A, BC and D), as shown in Figure 1b, with each of their cells having a transverse segmentation of $\Delta\eta \times \Delta\phi = 0.1 \times 0.1$ (0.2×0.1 in the outer D layer). Together with the electromagnetic calorimeter, TileCal provides precise measurement of hadrons, jets, taus and missing transverse energy.

3. Granularity Improvement Method

Aiming at extracting additional information on the spatial distribution of the energy deposited within each cell, the original TileCal PMT, Hamamatsu model R7877, may be substituted by a Multi-Anode Photomultiplier Tube (MA-PMT), Hamamatsu model R7600-300-M64, which has 64 photo sensors distributed in a grid of 8×8 pixels, in this particular case. To make detector granularity improvement feasible, an algorithm has to be developed to match a given image pattern formed in the grid of pixels to a topological sub region within a given cell.

As mentioned in Section 1, jets with high transverse momentum tend to deposit energy deeper into TileCal. For a jet with $p_T > 3$ TeV, on average, 50% of energy sampled by TileCal will be deposited within the BC layer. Therefore, splitting the BC layer into two separate layers, B and C, would help physicists to gain more information on the longitudinal shower profile of particles, and improve angular position measurements and jet energy resolution.

3.1. Calibration Data

To create the target mapping from a particular image pattern collected by the MA-PMT grid of pixels, we will use data from the Cesium calibration system taken during the Cs scan runs [8]. The Cs source system consists of a radioactive source of Cesium-137, which moves through the calorimeter by means of a hydraulic system, in order to calibrate the optical components of TileCal.

During the Cs scan, the readout electronics acquire simultaneously integrated currents from one channel connected to the single anode PMT and 48 channels connected to the new MA-PMT readout. In other words, both PMTs, single and multi-anode, are connected to the same calorimeter cells from different sides. Therefore, as the Cs source moves along the cell, performing sequential excitation of every tile, one can order the time-dependent response curves of each individual tile with respect to the images produced by the MA-PMT grid of pixels.

To identify images generated from B and C cells, we used the information about the geometry of the detector alongside with the time response curves produced during the Cs scan. Experimental data acquired comprises 1,647 images from the B (sub-cell) and 1,951 entries for the C (sub-cell). Despite being time consuming and demanding a significant amount of manpower, the Cs scan through TileCal cells results in a limited number of images, which makes deep learning development difficult to be accomplished. Recent developments in deep generative adversarial models, such as [4] and [5], could solve this issue.

3.2. Generative Adversarial Networks (GANs)

The generative model will be developed for increasing the statistics. The GAN approach is based on two deep neural networks that compete with each other. The first network is called the Generator, and its objective is to learn the mapping $\mathbf{x} = G(\mathbf{z}; \theta_g)$, over the data \mathbf{x} , where θ_g are the network parameters and \mathbf{z} is an input vector containing just noise. The second network is called the Discriminator and has a mapping $D(\mathbf{x}; \theta_d)$, which represents the probability of the data \mathbf{x} being real or generated by the G network. The Discriminator is trained to minimize the following cross-entropy cost function, where E_x and E_z are the expected values for the vectors \mathbf{x} and \mathbf{z} , respectively:

$$J^D = -\frac{1}{2}E_{x \sim p_{data}} \log D(x) - \frac{1}{2}E_z \log(1 - D(G(z))).$$

We found that the primary DCGAN architecture, used by the original authors, performs well in the MA-PMT images with a few modifications on the size of filters and kernels. These network parameters had to be modified to match the size of our 8×8 images. In the Generator network, we have three convolutional layers with 128, 64 and 1 filters, respectively. All the convolutions

use symmetric kernels with size 3 and strides 1. The Rectified Linear Unit (ReLU) [9] activation is used for all except the output layer, which uses a hyperbolic-tangent function.

The Discriminator has four convolutional layers, activated with LeakyReLU [10] with parameter $\alpha = 0.2$, and one dense layer as output with a sigmoid function. The first two convolutional layers have 32 and 64 filters, respectively, with symmetric kernels of size 3 and strides 2. The last two convolutional layers have 128 and 256 filters, with symmetric kernels with size 3 and strides 1. To train the adversarial network, the Adam [11] optimizer was used, with learning rate 0.0002 and $\beta_1 = 0.5$, as suggested by the original paper. To avoid model instability, batch normalization is used for all except the first layer of the Discriminator and the output layer of the Generator model.

4. Results

The DCGAN was trained for each sub-cell, B and C, separately. After the training process, the Generator model was used to produce 100,000 synthetic images used for the analysis.

4.1. Similarity Measure

After the generation process, all the images, real or synthetic, were converted to 8-bit grayscale images, with pixel intensity in the interval $[0, 255]$, in order to be compared. Figure 2 shows the average pixel intensity for both real and synthetic images, alongside with its relative error, compared pixel per pixel. The largest difference from the average images is around 11% for the B sub-cell and 14% for the C sub-cell.

The relative differences between real and synthetic images (Figure 3) also show that the GAN was able to reproduce the discriminating pattern between the B and C sub-cells.

4.2. Sub-Cell Classification

For the sub-cell classification, we trained the ConvNet using the synthetic data and test it with the real images obtained during the Cs scan data acquisition which were not used for training the generator model, in order to make the test set statistically independent from the training set.

The ConvNet architecture has only one convolution layer, with max pooling operation, followed by dense layers and an output layer. The activation function was ReLU for all layers, with exception of the output layer which has a sigmoid function. The ConvNet was trained with the Adam optimizer for 5 epochs.

A training accuracy of 0.9886 ± 0.0027 was obtained for the training set (synthetic images), which was slightly reduced to 0.9861 for the test set (real images). Table 1 shows the classification accuracy results. The model was able to correctly predict all the images corresponding to the B sub-cell and missclassified only 2.5% of the C sub-cell.

Table 1: Accuracy for the test set.

	Predicted B sub-cell	Predicted C sub-cell
True B sub-cell	1647	0
True C sub-cell	50	1901

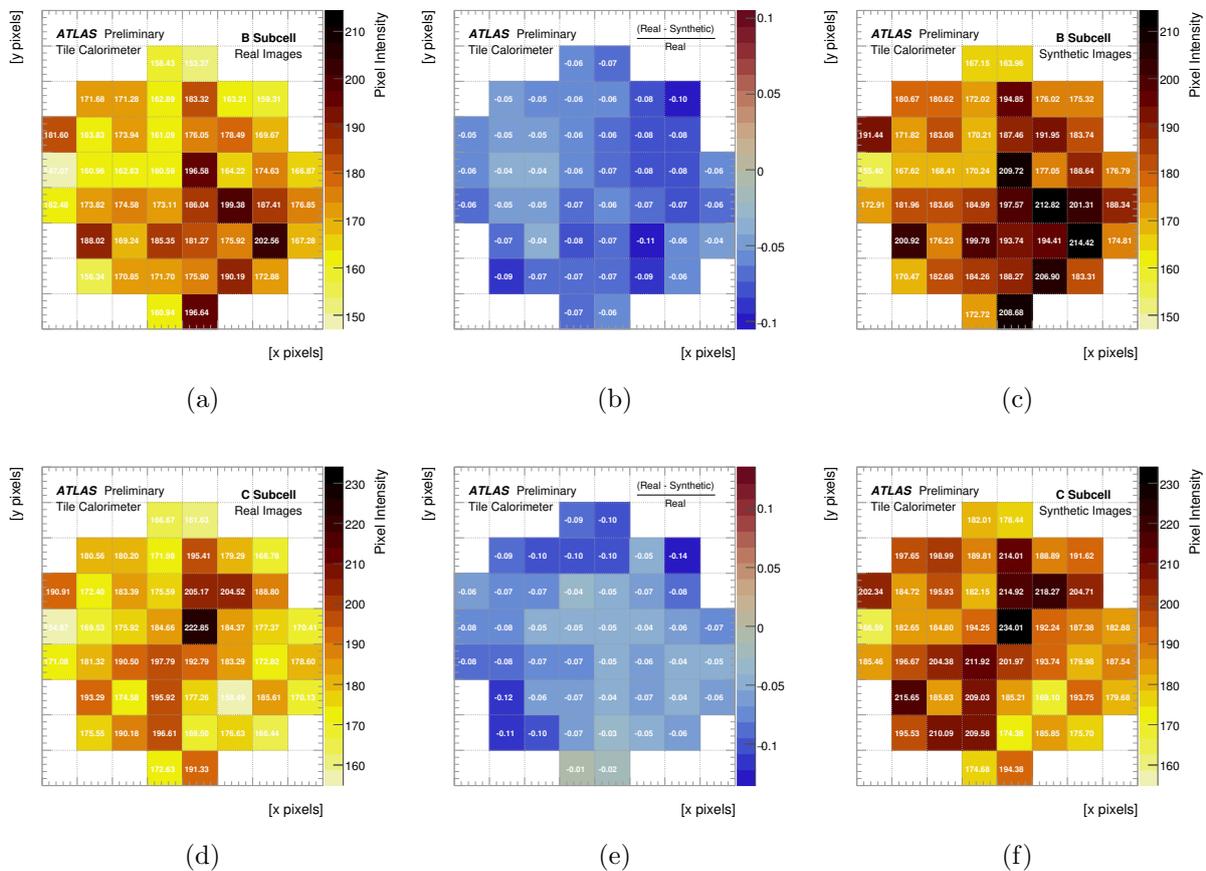


Figure 2: Average [pixel intensities] for real (left) and synthetic (right) images alongside with its relative difference (center). The first row shows the result for the B sub-cell, while the second row shows the results for the C sub-cell.

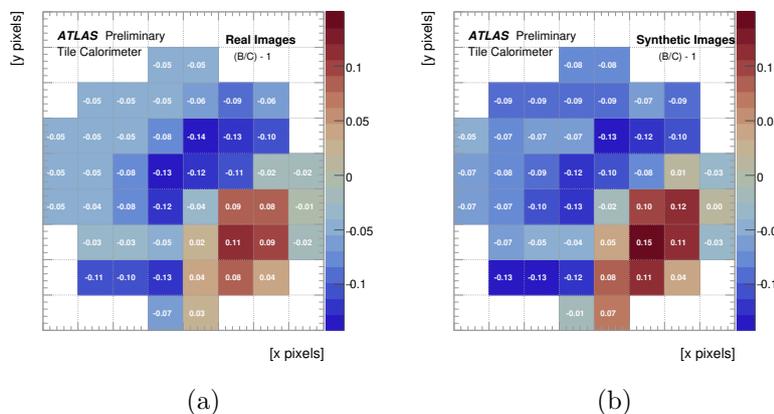


Figure 3: Relative differences (discriminating pattern) between B and C sub-cells for real (left) and synthetic (right) images.

5. Conclusions

Results of the recent efforts, based on the R&D studies to improve the TileCal granularity for the high luminosity regime of the LHC (HL-LHC) were shown. After splitting the BC cell into two separate sub-cells, we used deep learning methods to tackle the problem from a supervised perspective. First, increasing the statistics with simulation generated by an adversarial model, and then performing a binary classification algorithm based on deep neural networks.

The combination of the techniques, GANs and ConvNets, made it possible to look at the improvement of granularity in TileCal as a supervised learning problem. The first results of the binary classification showed a 98% accuracy in the test set, an improvement of around 26 percentage points when compared with a CNN trained only in real data. This method showed the possibility of increasing the actual longitudinal granularity of the BC cell by a factor of two, without changing the mechanical structure of the detector.

Acknowledgements

The following Brazilian funding agencies have supported this work: CNPq, CAPES, FAPERJ and RENAFAP (MCTIC).

References

- [1] Aad G *et al.* 2008 *J. of. Inst.* **03** S08003
- [2] Aad G *et al.* 2010 *Eur. Phys. J. C* **70** 1193–1236
- [3] Goodfellow I, Pouget-Abadie J, Mirza M *et al.* 2014 Generative adversarial networks (*Preprint 1406.2661*)
- [4] Paganini M, Oliveira L and Nachman B 2018 *Phys. Rev D* **97** 014021
- [5] Mustafa M, Bardi D, Bhimji W *et al.* 2019 *Comp. Astophys. and Cosm.* **06** 19–29
- [6] Radford A, Metz L and Chintala S 2015 Unsupervised representation learning with deep convolutional generative adversarial networks (*Preprint 1511.06434*)
- [7] Lecun Y, Bottou L, Bengio Y and Haffner P 1998 *Proc. IEEE* **86** 2278–324
- [8] Marjanovic M 2019 *IEEE Trans. on Nuclear Science* **66** 1228–35
- [9] Nair V and Hinton G 2010 *Proc. Int. Conf. on Machine Learning* **27** 807–14
- [10] Mass A L, Hannun A and Ng A Y 2013 *Proc. Int. Conf. on Machine Learning* **30** 03
- [11] Kingma D P and Ba J 2015 Adam: A method for stochastic optimization (*Preprint 1412.6980*)