



Original software publication

# HexagDly—Processing hexagonally sampled data with CNNs in PyTorch

Constantin Steppa <sup>a,\*</sup>, Tim L. Holch <sup>b,1</sup><sup>a</sup> Experimental Astroparticle Physics, Department of Physics and Astronomy, University of Potsdam, Karl-Liebknecht-Straße 24/25, 14476 Potsdam-Golm, Germany<sup>b</sup> Experimental Particle Physics, Department of Physics, Humboldt University of Berlin, Newtonstr. 15, 12489 Berlin, Germany

## ARTICLE INFO

## Article history:

Received 26 November 2018

Received in revised form 28 February 2019

Accepted 28 February 2019

## Keywords:

Convolutional neural networks

Hexagonal grid

PyTorch

Astroparticle physics

## ABSTRACT

HexagDly is a Python-library extending the PyTorch deep learning framework with convolution and pooling operations on hexagonal grids. It aims to ease the access to convolutional neural networks for applications that rely on hexagonally sampled data as, for example, commonly found in ground-based astroparticle physics experiments.

© 2019 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

## Code metadata

Current code version  
 Permanent link to code/repository used for this code version  
 Legal Code License  
 Code versioning system used  
 Software code languages, tools, and services used  
 Compilation requirements, operating environments & dependencies  
 If available Link to developer documentation/manual  
 Support email for questions

2.0.1  
[https://github.com/ElsevierSoftwareX/SOFTX\\_2018\\_240](https://github.com/ElsevierSoftwareX/SOFTX_2018_240)  
 MIT  
 git  
 Python 3, PyTorch  
 Tested on Linux and Mac OS, Python  $\geq 3.6$ , PyTorch  $\geq 0.4$   
<https://github.com/ai4iacts/hexagdly>  
[steppa@uni-potsdam.de](mailto:steppa@uni-potsdam.de), [holchtim@physik.hu-berlin.de](mailto:holchtim@physik.hu-berlin.de)

## 1. Motivation and significance

Convolutional neural networks (CNNs) are a powerful and versatile tool in big data analysis and computer vision [1]. Their application has been widely promoted in various research fields by the availability of open-source deep learning frameworks (DLFs) like TensorFlow, Caffe, PyTorch or the Microsoft Cognitive Toolkit. Also in ground-based astroparticle physics experiments, where large amounts of image-like data need to be analysed, the application of CNNs has come into focus.

This data is often hexagonally sampled, which poses an initial obstacle for the application of CNNs: DLFs cannot process hexagonally sampled data out-of-the-box. Solutions to this problem

have been presented in several applicability studies [2–7]. Most of these solutions are based on transforming the hexagonally sampled data to an approximate representation on a rectangular grid via pre-processing such as rebinning, interpolation, oversampling and axis-shearing. HexagDly, on the other hand, provides a native solution to process hexagonally sampled data. It relies on a specific addressing scheme for hexagonally sampled data that allows for the construction of convolution and pooling operations on hexagonal grids by using methods provided by PyTorch<sup>2</sup> [8]. HexagDly thereby aims to exploit the benefits of directly processing hexagonally sampled data, of which the most notable are reduced computing resources [9], more efficient image processing operators [10] and higher angular resolution [11].

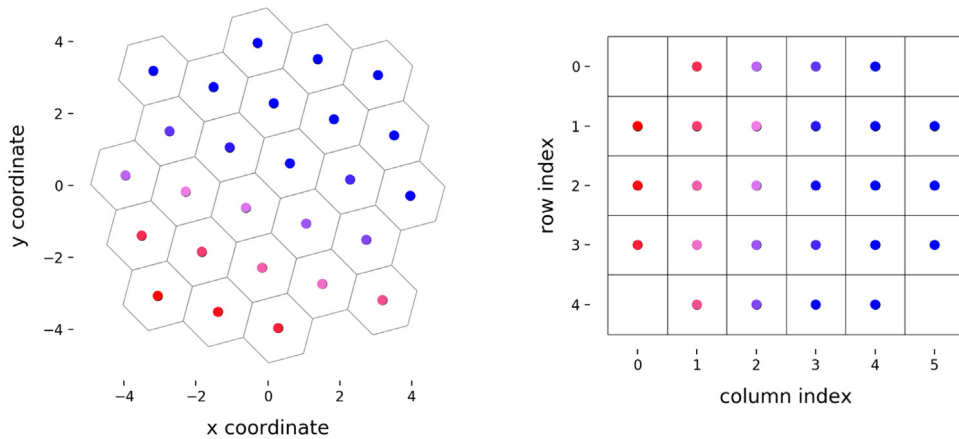
In the context of CNNs, Hooeboom et al. have already demonstrated the advantages of applying hexagonal convolutions such

\* Corresponding author.

E-mail addresses: [steppa@uni-potsdam.de](mailto:steppa@uni-potsdam.de) (C. Steppa), [holchtim@physik.hu-berlin.de](mailto:holchtim@physik.hu-berlin.de) (T.L. Holch).

<sup>1</sup> Both authors contributed equally.

<sup>2</sup> <https://pytorch.org/>.



**Fig. 1.** Illustrative example of a hexagonal array with pixel positions given in Cartesian coordinates (left) and the corresponding tensor indices as inferred from the described addressing scheme (right). Blank elements of the tensor have to be filled with arbitrary values.

as improved accuracies due to the reduced anisotropy of hexagonal filters [12]. With HexagDLY, hexagonal convolutions are available in an open-source software with focus on user-friendliness. It facilitates access to CNNs for any kind of hexagonally sampled data, which, in addition to ground-based astroparticle physics, can be found in other research fields like ecology [13] or numerical climate modelling [14,15].

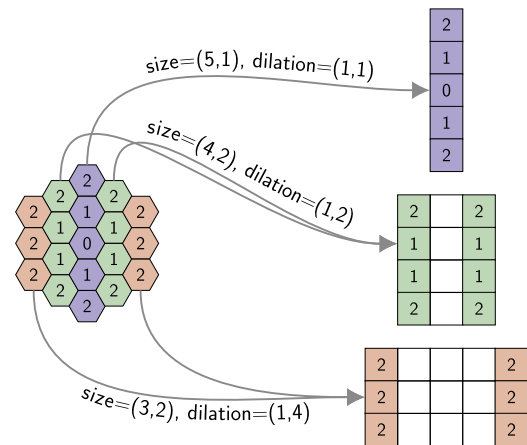
In the following Section 2 the software is described, including its capabilities and the requirements on the input format. The application of HexagDLY is illustrated with an example in Section 3 followed by a comparative study on the application of hexagonal and square convolution kernels. Potential benefits of using hexagonal convolutions in ground-based astroparticle physics are outlined in Section 5.

## 2. Software description

HexagDLY provides convolution operations on hexagonal grids built on PyTorch routines. Given the required input format for these routines, an addressing scheme has to be chosen to map the hexagonally sampled data to Cartesian tensors. The convolution and pooling operations are then adapted accordingly to reflect the hexagonal structure of the original data which is also conserved in the output. This is done by constructing custom hexagonal kernels that are applied in combination with a strict padding and striding scheme. These main ideas behind HexagDLY are outlined below. Please see Code metadata for the repository and software dependencies.

### 2.1. Input format

In order to map hexagonally sampled data to Cartesian tensors, different addressing schemes can be applied (for example, see [12]). HexagDLY uses the scheme that allows for the most efficient data storage. As a hexagonal grid can be interpreted as two overlaid rectangular grids, the data points can be combined in a single square-grid array by aligning the two rectangular parts. The procedure is illustrated in Fig. 1 where the hexagonal array in Cartesian coordinates is first rotated to achieve a vertical alignment of neighbouring elements (called *pixels* hereafter). This allows a separation of the data into columns. The pixels are then aligned horizontally by shifting every second column upwards by half the distance between neighbouring pixels, resulting in a square-grid array with rows and columns. Counting rows from top to bottom and columns from left to right yields the indices for each element in the input tensor, which corresponds to a certain pixel in the hexagonal array. Tensor elements that do not have a corresponding counterpart in the hexagonal array have to be filled with an arbitrary value.



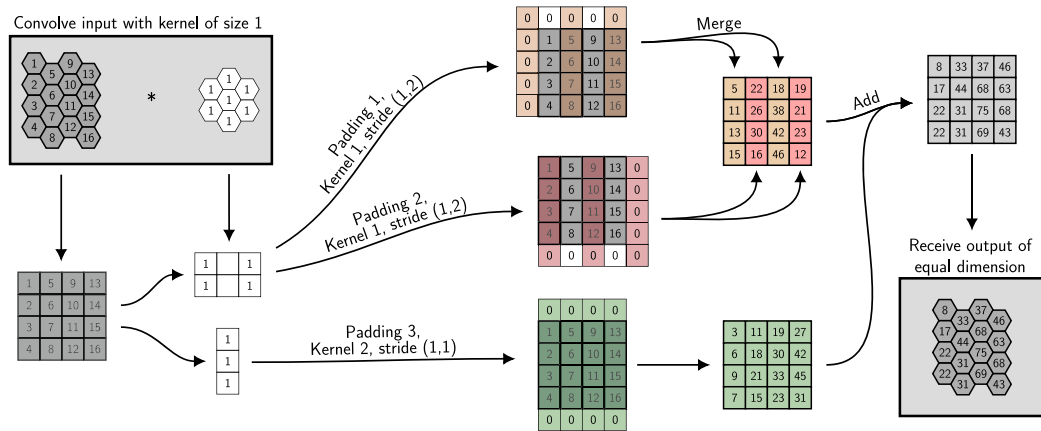
**Fig. 2.** Schematic construction of a hexagonal kernel from rectangular sub-kernels within HexagDLY. Only the exterior columns of sub-kernels contain values while interior columns are disregarded by setting a *dilation*  $\neq 1$  (see the PyTorch documentation for details).

### 2.2. Hexagonal kernels

The implemented convolution operations use kernels on the hexagonal grid that have a 6-fold rotational symmetry (i.e. kernels of hexagonal shape). The geometry of a kernel is therefore described only by its size which is a single integer corresponding to the number of layers of neighbouring elements around its central element. Internally, HexagDLY constructs these hexagonal kernels from rectangular sub-kernels as illustrated in Fig. 2. The illustrated kernel of size 2 consists of three sub-kernels, each representing a set of equal-length columns of the hexagonal kernel. The spatial relation between these columns are accounted for via defined horizontal dilations.

### 2.3. Convolution operations

Since a hexagonal kernel is constructed out of multiple rectangular sub-kernels, a single hexagonal convolution operation is realised by a combination of multiple convolutions of the input tensor with these sub-kernels. As described in Section 2.1, columns of the hexagonal array are shifted to match with the tensor format required by PyTorch. The single sub-convolutions therefore have to be adapted in order to account for this shift. This is achieved by defining a complex scheme for the padding



**Fig. 3.** Realisation of a hexagonal convolution with a kernel of size 1 in HexagDly. First, the input data is rearranged into a tensor (as described in Section 2.1) and the kernel is divided into rectangular sub-kernels. For every sub-kernel, different paddings and strides are applied to the input to account for the shifted columns. The results of the sub-convolutions are then merged and added to receive the convolved hexagonal data in tensor format.

and slicing of the input tensor. In this scheme, the number of rows and columns that are padded or sliced for each sub-convolution depends on the size of the input tensor as well as on the size of the hexagonal kernel and the applied stride. To conserve the hexagonal structure of the data, only symmetric strides in equally sized steps along the three symmetry axes of the hexagonal grid are performed, starting from the top left cell. Fig. 3 illustrates the single steps of this procedure for the convolution of a toy tensor with a hexagonal kernel of size 1.

It is important to note that a kernel is always centred on a pixel that is part of the actual input tensor and not of the padded rows and columns. To conserve the data format used by HexagDly, steps that would lead to an output with columns of unequal length are neglected. Fig. 4 illustrates this padding and convolution-element selection for different strides and kernel sizes, including such a case, where a step is omitted.

#### 2.4. Software functionalities

HexagDly provides two- and three-dimensional hexagonal convolution operations. In the three-dimensional case, the input data is expected to have a hexagonal layout in the x-y-plane while data points along the z-axis are assumed to be equidistant. This makes it possible e.g. to process time-resolved data of a two-dimensional detector with hexagonal layout. Following the design of convolution operations, pooling methods are implemented accordingly. This is done by replacing the PyTorch-based sub-convolutions with the according pooling methods and combining the outputs with aggregation functions, whereas the padding- and striding-scheme is identical. By adopting the PyTorch-API, these operations can easily be incorporated in CNN models defined in PyTorch. Furthermore, it is possible to define custom hexagonal kernels with defined values for each kernel element, making it possible to manually implement structure detecting kernels or to perform data processing like smoothing on hexagonally sampled data. Examples are provided in the online repository in the form of jupyter notebooks (see Code metadata) that demonstrate the functionalities and usage of the methods provided by HexagDly.

### 3. Illustrative example

To outline the application of HexagDly, a set of examples covering basic use-cases is provided along with the HexagDly source code in the online repository. An illustrative way to demonstrate

the functioning and capabilities of HexagDly is to perform hexagonal operations on hexagonally sampled shapes that themselves exhibit a 6-fold symmetry. In Fig. 5 the result of convolving an image displaying hexagonal shapes with a hexagonal kernel is shown. It can clearly be seen that the 6-fold symmetry of the original shapes on the hexagonal grid is conserved in the output.

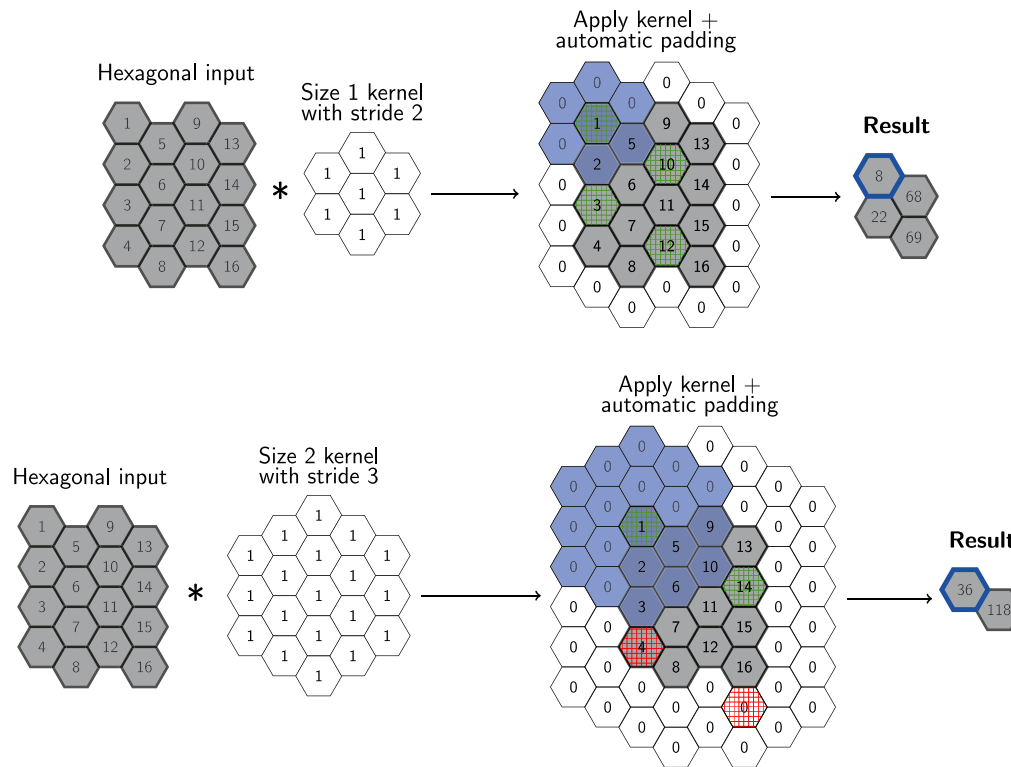
For an example of how to use HexagDly in a CNN, please see the provided jupyter notebooks in the online repository (see Code metadata).

### 4. Comparing hexagonal and square convolution kernels

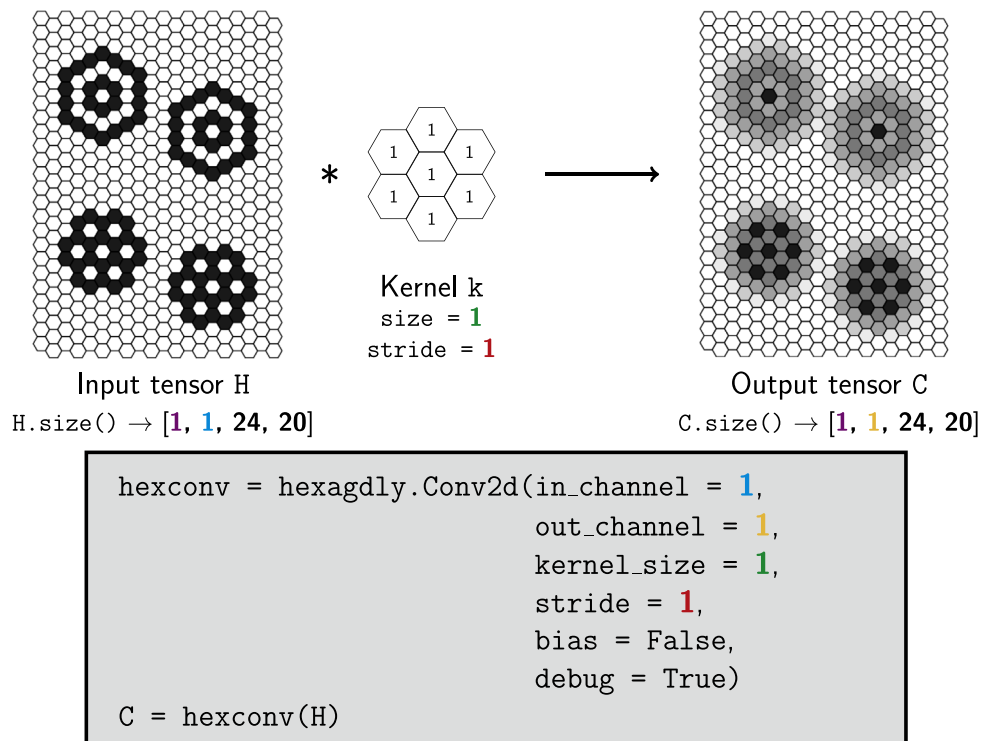
As outlined in Section 1, a hexagonal sampling of two-dimensional data allows for more efficient data processing compared to a square-grid sampling. Starting with hexagonally sampled data, a conversion to a square grid representation therefore implies less efficient data processing. Additionally, re-sampling hexagonally sampled data to a square grid can introduce sampling artefacts and often requires an increase in resolution to reduce distortions.

In the context of deep learning, the effects of re-sampling and the reduction of processing efficiency can have a significant influence on the process of designing, optimising and applying CNN-based algorithms. While the applied re-sampling method is an independent parameter that can be optimised, an increase in resolution demands more computer storage and implies larger convolution kernels or more convolution layers to retain a certain receptive field. In combination, these effects can influence the performance of a CNN significantly. This is demonstrated in the following by comparing the performance of CNNs that are trained for the same task but use either hexagonal- or square-grid operations on hexagonal or re-sampled data, respectively.

For the presented experiment, a data set was created with images of four different hexagonal shapes at random positions on a hexagonal grid, overlaid with Gaussian noise. This data set was then interpolated to a square grid of the same resolution (small) as well as to a square grid with four times the number of pixels (large). An example of such a hexagonal shape with the according re-sampled images is shown in Fig. 6. Two CNN models with the same architecture were set up with the only difference being the use of hexagonal (*h-CNN, small*) or square-grid operations (*s-CNN, small*). These two models have two convolutional and three fully connected layers with a total of ~13k learnable parameters. A third CNN model with three convolutional and three fully connected layers and a total of ~1.2M learnable parameters (*s-CNN, large*) was set up and trained on the large square-grid data.



**Fig. 4.** Illustration of the padding and striding scheme for convolutions with different kernel sizes and strides. The green and red elements mark valid and omitted steps, respectively. The first position of a kernel is marked in blue as well as its corresponding output cell in the result.



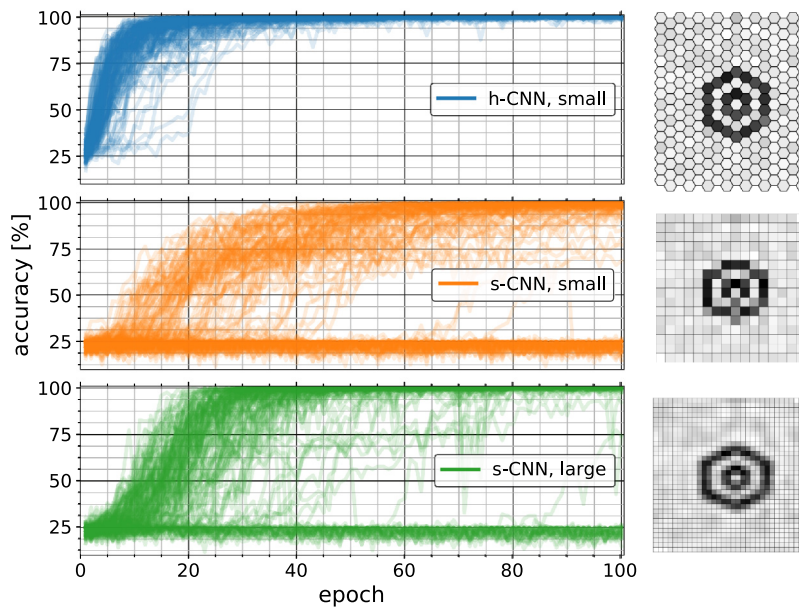
**Fig. 5.** Schematic application of a hexagonal kernel of size 1 to hexagonal shapes on a hexagonal grid. The corresponding code is given in the grey box, whereas the parameters defining the operation are colour-coded. Enabling the debug mode sets all kernel elements to 1.

The full implementation of the CNN-models and the data set are provided in a jupyter notebook in the online repository. The three CNNs were trained for 100 epochs on 128 images per class with a self-adjusting learning rate. This was repeated 150 times with the

training data being regenerated and the models being reinitialised in each iteration.

Fig. 6 shows the resulting learning curves for all iterations for each CNN-model. It can be seen that the h-CNN reliably reaches





**Fig. 6.** Learning curves of 150 iterations for the three CNN models trained to distinguish between four different hexagonal shapes. Example images of one of the four shapes are shown in their different samplings right of each learning curve. See Section 4 for details on the CNN models and data sets.

100% accuracy after a few epochs of training. Both s-CNNs, on the other hand, show a generally worse learning behaviour. Although they are both able to achieve 100% accuracy in some cases, only in 60% (small) and 80% (large) of all iterations the models reach accuracies above random guessing performance.

This toy example illustrates the advantages of directly processing hexagonally sampled data in terms of reliability and accuracy. The differences in performance of the two s-CNNs demonstrate that the effects of re-sampling can be compensated by increasing the resolution of the re-sampled data and likewise extending the CNN capacity. However, even with two orders of magnitude more learnable parameters, the performance of the h-CNN is not reached. Even though the performance difference between h-CNN and s-CNN may not be as significant in a realistic application, natively processing hexagonally sampled data is generally expected to be the most efficient approach. However, the current implementation of hexagonal operations in HexagDLY produces a significant computational overhead compared to its according square-grid operation in PyTorch. This can increase the processing time for an h-CNN implemented with HexagDLY, but does not influence the advantages of applying hexagonal convolutions as outlined above.

## 5. Impact

Hexagonally sampled data is common in ground-based astroparticle physics experiments like the High Energy Stereoscopic System (H.E.S.S.), the Pierre Auger Observatory or IceCube where large areas have to be efficiently covered with a limited number of detectors. This can be achieved by arranging the detectors on a hexagonal grid as it allows for the densest tiling of a two-dimensional Euclidean plane and for optimal sampling of circularly band-limited signals. In these experiments data is taken at high rates and is mostly background-dominated. Additionally, this data can cover a large parameter space, e.g. multiple telescopes taking data simultaneously. Therefore, advanced data processing algorithms are used to analyse this data. The application of machine learning techniques has already become a standard in this respect [16,17]. Following the progress in the field of machine learning, CNNs represent promising means to further improve data analyses for astroparticle physics experiments.

By providing convolution and pooling operations that can be directly applied to hexagonally sampled data, HexagDLY provides a user-friendly environment to explore the applicability of CNNs for these experiments. Since no pre-processing is required, the initial efforts for the application of CNNs can be significantly reduced compared to other approaches.

The increasing scales and sensitivity of future observatories like the Cherenkov Telescope Array [18] will result in much larger data sets that need to be analysed. This will pose additional challenges for the analysis in terms of performance and resources. The methods provided by HexagDLY can help to address these challenges.

## 6. Conclusions

Following the growing interest in CNNs, increasing efforts to adapt convolution operations to non-Cartesian data can be observed, as for example for spherical data [19,20] and non-Euclidean manifolds [21]. Besides [12], HexagDLY presents a solution for hexagonally sampled data. With a focus on flexibility and user-friendliness, HexagDLY provides convolution and pooling operations on hexagonal grids. It is based on PyTorch and makes use of the torch.nn module for the implementation of these operations. In combination with a special data addressing scheme, it facilitates the access to CNNs for hexagonally sampled data. By taking advantage of the benefits of directly processing hexagonally sampled data, HexagDLY aims to promote research based on the applicability of CNNs e.g. in ground-based astroparticle physics. Currently, HexagDLY is used in a study on the applicability of CNNs for the analysis of data from the H.E.S.S. experiment. A report on first results is in preparation.

## Acknowledgements

This software project evolved as part of a research study to explore deep learning algorithms in the analysis of imaging Cherenkov telescope data within the H.E.S.S. collaboration. We acknowledge the support of the whole collaboration in this project. In particular we want to thank our colleagues Matthias Büchele, Kathrin Egberts, Tobias Fischer, Manuel Kraus, Thomas Lohse, Ullrich Schwanke, Idan Shilon and Gerrit Spengler for fruitful discussions that promoted the development of HexagDLY.

## References

- [1] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436–44. <http://dx.doi.org/10.1038/nature14539>.
- [2] Feng Q, Lin TTY. The analysis of VERITAS muon images using convolutional neural networks, in: *Proceedings of the International Astronomical Union*, vol. 12, 2016.
- [3] Holch TL, Shilon I, Büchele M, Fischer T, Funk S, Groeger N, Jankowsky D, Lohse T, Schwanke U, Wagner P. Probing convolutional neural networks for event reconstruction in  $\gamma$ -ray astronomy with Cherenkov telescopes, in: *PoS ICRC2017, The Fluorescence detector Array of Single-pixel Telescopes: Contributions to the 35th International Cosmic Ray Conference (ICRC 2017)*, p. 795, <http://dx.doi.org/10.22323/1.301.0795>, arXiv:1711.06298.
- [4] Huennefeld M. Deep learning in physics exemplified by the reconstruction of muon-neutrino events in IceCube, in: *PoS ICRC2017, The Fluorescence detector Array of Single-pixel Telescopes: Contributions to the 35th International Cosmic Ray Conference (ICRC 2017)*, p. 1057, <http://dx.doi.org/10.22323/1.301.1057>.
- [5] Erdmann M, Glombitza J, Walz D. A deep learning-based reconstruction of cosmic ray-induced air showers. *Astropart Phys* 2018;97:46–53. <http://dx.doi.org/10.1016/j.astropartphys.2017.10.006>, URL <http://www.sciencedirect.com/science/article/pii/S0927650517302219>.
- [6] Mangano S, Delgado C, Bernardos M, Lallena M, Vázquez JJR. Extracting gamma-ray information from images with convolutional neural network methods on simulated cherenkov telescope array data. In: *ANNPR 2018. LNAI*, vol. 11081, 2018, p. 243–54. <http://dx.doi.org/10.1007/978-3-319-99978-4>, arXiv:1810.00592.
- [7] Shilon I, Kraus M, Büchele M, Egberts K, Fischer T, Holch TL, Lohse T, Schwanke U, Steppa C, Funk S. Application of deep learning methods to analysis of imaging atmospheric cherenkov telescopes data. *Astropart Phys* 2019;105:44–53. <http://dx.doi.org/10.1016/j.astropartphys.2018.10.003>, URL <http://www.sciencedirect.com/science/article/pii/S0927650518301178>.
- [8] Paszke A, Gross S, Chintala S, Chanan G, Yang E, DeVito Z, Lin Z, Desmaison A, Antiga L, Lerer A. Automatic differentiation in PyTorch, in: *NIPS-W*, 2017.
- [9] Mersereau RM. The processing of hexagonally sampled two-dimensional signals. *Proc IEEE* 1979;67(6):930–49. <http://dx.doi.org/10.1109/PROC.1979.11356>.
- [10] Staunton R. The design of hexagonal sampling structures for image digitization and their use with local operators. *Image Vis Comput* 1989;7(3):162–6. [http://dx.doi.org/10.1016/0262-8856\(89\)90040-1](http://dx.doi.org/10.1016/0262-8856(89)90040-1).
- [11] Richard C. Staunton NS. A comparison between square and hexagonal sampling methods for pipeline image processing. *Proc SPIE* 1990;1194:1194 – 1194 – 10.
- [12] Hoogetboom E, Peters JWT, Cohen TS, Welling M. Hexaconv. 2018, ArXiv e-prints [arXiv:1803.02108](https://arxiv.org/abs/1803.02108).
- [13] Birch C, P. Oom S, A. Beecham J. Rectangular and hexagonal grids used for observation, experiment and simulation in ecology. *Ecol Model* 2007;206:347–59. <http://dx.doi.org/10.1016/j.ecolmodel.2007.03.041>.
- [14] Sahr K. Hexagonal discrete global grid systems for geospatial computing. *Arch Photogramm Cartogr Remote Sens* 2011;22:363–76.
- [15] Satoh M, Tomita H, Yashiro H, Miura H, Kodama C, Seiki T, Noda AT, Yamada Y, Goto D, Sawada M, Miyoshi T, Niwa Y, Hara M, Ohno T, Iga S-i, Arakawa T, Inoue T, Kubokawa H. The non-hydrostatic icosahedral atmospheric model: description and development. *Prog Earth Planet Sci* 2014;1(1):18. <http://dx.doi.org/10.1186/s40645-014-0018-1>.
- [16] Ohm S, van Eldik C, Egberts K.  $\gamma$ /hadron separation in very-high-energy  $\gamma$ -ray astronomy using a multivariate analysis method. *Astropart Phys* 2009;31:383–91. <http://dx.doi.org/10.1016/j.astropartphys.2009.04.001>, arXiv:0904.1136.
- [17] Aartsen et al. MG. Search for dark matter annihilation in the galactic center with icecube-79. *Eur Phys J C* 2015;75:492. <http://dx.doi.org/10.1140/epjc/s10052-015-3713-1>.
- [18] T. Cherenkov Telescope Array Consortium, B.S. Acharya, I. Agudo, I.A. Samarai, R. Alfaro, J. Alfaro, C. Alispach, R. Alves Batista, J.-P. Amans, et al., Science with the Cherenkov Telescope Array, 2017, ArXiv e-prints, [arXiv:1709.07997](https://arxiv.org/abs/1709.07997).
- [19] Cohen TS, Geiger M, Khler J, Welling M. Spherical CNNs, in: *International Conference on Learning Representations*, 2018, URL <https://openreview.net/forum?id=Hkdb5xZRB>.
- [20] Perraudin N, Defferrard M, Kacprzak T, Sgier R. DeepSphere: Efficient spherical convolutional neural network with HEALPix sampling for cosmological applications. 2018, arXiv:1810.12186.
- [21] Masci J, Boscaini D, Bronstein MM, Vandergheynst P. Shapenet: convolutional neural networks on non-euclidean manifolds. 2015, URL <http://arxiv.org/abs/1501.06297>.