

Deep Learning in Liquid argon time projection chambers: Optimizing the Architecture for Particle Classification

MicroBooNE Collaboration

R. Acciarri^g C. Adams^z R. An^h J. Asaadi^w M. Auger^a L. Bagby^g B. Baller^g G. Barr^q M. Bass^q
F. Bay^x M. Bishai^b A. Blake^j T. Boltonⁱ L. Bugel^m L. Camilleri^f D. Caratelli^f B. Carls^g
R. Castillo Fernandez^g F. Cavanna^g H. Chen^b E. Church^r D. Cianci^{l,f} G. H. Collin^m
J. M. Conrad^m M. Convery^u J. I. Crespo-Anadón^f M. Del Tutto^q D. Devitt^j S. Dytman^s
B. Eberly^u A. Ereditato^a L. Escudero Sanchez^c J. Esquivel^v B. T. Fleming^z W. Foreman^d
A. P. Furmanski^l G. T. Garvey^k V. Genty^f D. Goeldi^a S. Gollapinniⁱ N. Graf^s E. Gramellini^z
H. Greenlee^g R. Grosso^e R. Guenette^q A. Hackenburg^z P. Hamilton^v O. Hen^m J. Hewes^l
C. Hill^l J. Ho^d G. Horton-Smithⁱ C. James^g J. Jan de Vries^c C.-M. Jen^y L. Jiang^s
R. A. Johnson^e B. J. P. Jones^m J. Joshi^b H. Jostlein^g D. Kaleko^f G. Karagiorgi^{l,f}
W. Ketchum^g B. Kirby^b M. Kirby^g T. Kobilarcik^g I. Kreslo^a A. Laube^q Y. Li^b A. Lister^j
B. R. Littlejohn^h S. Lockwitz^g D. Lorca^a W. C. Louis^k M. Luethi^a B. Lundberg^g X. Luo^z
A. Marchionni^g C. Mariani^y J. Marshall^c D. A. Martinez Caicedo^h V. Meddageⁱ T. Miceli^o
G. B. Mills^k J. Moon^m M. Mooney^b C. D. Moore^g J. Mousseauⁿ R. Murrells^l D. Naples^s
P. Nienaber^t J. Nowak^j O. Palamara^g V. Paolone^s V. Papavassiliou^o S.F. Pate^o Z. Pavlovic^g
D. Porzio^l G. Pulliam^v X. Qian^b J. L. Raaf^g A. Rafiqueⁱ L. Rochester^u C. Rudolf von Rohr^a
B. Russell^z D. W. Schmitz^d A. Schukraft^g W. Seligman^f M. H. Shaevitz^f J. Sinclair^a
E. L. Snider^g M. Soderberg^v S. Söldner-Rembold^l S. R. Soleti^q P. Spentzouris^g J. Spitzⁿ
J. St. John^e T. Strauss^g A. M. Szelc^l N. Tagg^p K. Terao^f M. Thomson^c M. Touns^g Y.-T. Tsai^u
S. Tufanli^z T. Usher^u R. G. Van de Water^k B. Viren^b M. Weber^a J. Weston^c
D. A. Wickremasinghe^s S. Wolbers^g T. Wongjirad^m K. Woodruff^o T. Yang^g G. P. Zeller^g
J. Zennaro^d C. Zhang^b

^aUniversität Bern, Bern CH-3012, Switzerland

^bBrookhaven National Laboratory (BNL), Upton, NY, 11973, USA

^cUniversity of Cambridge, Cambridge CB3 0HE, United Kingdom

^dUniversity of Chicago, Chicago, IL, 60637, USA

^eUniversity of Cincinnati, Cincinnati, OH, 45221, USA

^fColumbia University, New York, NY, 10027, USA

^gFermi National Accelerator Laboratory (FNAL), Batavia, IL 60510, USA

^hIllinois Institute of Technology (IIT), Chicago, IL 60616, USA

ⁱKansas State University (KSU), Manhattan, KS, 66506, USA

^jLancaster University, Lancaster LA1 4YW, United Kingdom

^kLos Alamos National Laboratory (LANL), Los Alamos, NM, 87545, USA

^lThe University of Manchester, Manchester M13 9PL, United Kingdom

^mMassachusetts Institute of Technology (MIT), Cambridge, MA, 02139, USA

ⁿUniversity of Michigan, Ann Arbor, MI, 48109, USA

^oNew Mexico State University (NMSU), Las Cruces, NM, 88003, USA

^pOtterbein University, Westerville, OH, 43081, USA

^qUniversity of Oxford, Oxford OX1 3RH, United Kingdom

^rPacific Northwest National Laboratory (PNNL), Richland, WA, 99352, USA

^sUniversity of Pittsburgh, Pittsburgh, PA, 15260, USA

^tSaint Mary's University of Minnesota, Winona, MN, 55987, USA

^uSLAC National Accelerator Laboratory, Menlo Park, CA, 94025, USA

^vSyracuse University, Syracuse, NY, 13244, USA

^wUniversity of Texas, Arlington, TX, 76019, USA

^xTUBITAK Space Technologies Research Institute, METU Campus, TR-06800, Ankara, Turkey

^yCenter for Neutrino Physics, Virginia Tech, Blacksburg, VA, 24061, USA

^zYale University, New Haven, CT, 06520, USA

ABSTRACT: Liquid argon time projection chambers (LArTPCs) produce image-like data that may be analyzed with deep neural networks (Deep Learning, or DL). This approach to data analysis is becoming more common in LArTPCs, as well as other high channel-count segmented detectors where the events leave imprints on the detector that may be interpreted as images. In particular, MicroBooNE is an 8256 wire detector viewing 89 tons fiducial liquid argon where DL studies have been performed [2] and show great promise. That study showed progress in single particle classification as well as progress toward neutrino identification, using networks that were not necessarily chosen to be optimal. This paper reports results on single particle classification using DL network architectures in which the literature or experience suggest further progress can be made on images such as ours. We answer key questions about performance on this problem as a function of network architectures and show that networks __, __ and __ give functionally equivalent high performance, with gains in network __ in resource consumption that suggest its use.

Contents

1	Introduction	2
2	Particle Classification	2
3	Image Preparation	2
4	Training and Inference	3
5	The Architectures	3
5.1	Original CNN paper architectures	3
5.2	ResNet14b	3
5.3	ResNet14b_w2	3
5.4	vgg16a	3
5.5	vgg16b	3
5.6	vgg16c	3
6	Conclusions	3

1 Introduction

MicroBooNE is the first large-scale LArTPC detector to collect data in a neutrino beam in the United States. The experiment has a total liquid argon mass of 170 tons and a rectangular active volume of $2.3 \times 2.6 \times 10.4 \text{ m}^3$. The system consists of two subdetectors: a time projection chamber (TPC) for tracking and calorimetry along with a light collection system for the detection of scintillation light. We consider here only the TPC information. In October 2015, the detector began observing interactions of neutrinos from the FNAL Booster Neutrino Beam (BNB) [22].

In this paper We will survey the relevant networks for image detection that are typical for LArTPCs and compare performance metrics. A particular characteristic of our images that makes them unique compared to other efforts on image classification is that our images are sparse. Tracks and showers and isolated electromagnetic activity produce a relatively small fraction of pixels that are activated in LArTPCs. This is true even for MicroBooNE’s particular case in which the detector sits on the surface and sees a $\approx 5 \text{ kHz}$ of through-going cosmic rays and associated activity.

We want to survey networks that are well-used in the image recognition literature and explore the (hyper) parameters of those networks. This means we will explore the depth and width of networks, the existence of fully connected layers, stride sizes, as well as input sparsification and existence of things like “bottlenecks,” which is We will use variants of ResNet [], Vgg16 [] and ... One motivation of this exploration, in addition to seeking the best classification performance, is to see if networks that are less computationally expensive might in fact retain all performance of their more computationally expensive brethren networks, and thus might be used in their place.

The organization of this paper consists of a description of our single particle metrics, the preparation of the images, the training and inference protocol used here, and finally a description of the various architectures with results, followed by conclusions.

2 Particle Classification

We study the performance of five particle classification (e^- , γ , μ^- , π^- , proton) using CNNs.

3 Image Preparation

Throughout all studies in this work, we use one of the most popular open-source CNN software frameworks, CAFFE [11], for CNN training and analysis. Input data is in a ROOT file format [16] created by the LArCV software framework [10], which we developed to act as the interface between LArSoft and CAFFE. LArCV is also an image data processing framework and is used to further process and analyze images as described in the following sections. One can find our custom version of CAFFE that utilizes the LArCV framework in [10]. The computing hardware used in this study consists of multiple servers at various institutions. We have two dual-GPU 12 GByte memory, NVIDIA GPU servers of similar specifications at the Massachusetts Institute of Technology, Columbia University, University of Michigan, Yale University. We also employ a DGX-1 10-GPU 18 GB/GPU server at Pacific Northwest National Laboratories.

Seems like we ought to have a representative image picture here, as from the original CNN paper ...

Figure 1: Example neutrino candidate observed in the MicroBooNE detector. The same interaction is shown in all three wire plane views. The top image is from wire plane U ; the middle from wire plane V ; and the bottom from wire plane Y . The image is at full resolution and is only from a portion of the full event view.

4 Training and Inference

A disciplined training and inference procedure is required in order to make the comparisons we wish to make.

5 The Architectures

Broad descriptions here of ResNet and Vgg16 and whatever else we have results for. GoogleNet, AlexNet? Why we want to stick with flavors of ResNet and Vgg here.

Not to be included, but for convenience, all results areshould-be referenced at

<https://docs.google.com/presentation/d/1Rvo2Yt9KqRQk2eDy01WeA5GwaWamLvZR2CR3txfYbBE/edit#>

5.1 Original CNN paper architectures

5.2 ResNet14b

5.3 ResNet14b_w2

5.4 vgg16a

5.5 vgg16b

5.6 vgg16c

6 Conclusions

We have surveyed a variety of CNNs for the sake of classifying five charged particles as they appear in LArTPCs. Some interesting conclusions are that while Ref. [2] achieved formidable performance, a few other networks do as well or better, and it is seen now that RestNetXYZ and VGG16XYZ are better choices.

We plan to start with these networks for future physics-identification tasks in MicroBooNE; we believe similar guidance for general TPCs applies there, as well. We think, per this study, the full and infinite (hyper) parameter space for CNNs need not be explored in those future analyses.

References

- [1] Alex Krizhevsky *et al.*, “Imagenet Classification with Deep Convolutional Neural Networks,” *NIPS* **25**, pages 1106–1114, 2012.
- [2] R. Acciarri *et al.* [MicroBooNE collaboration], “Convolutional Neural Networks Applied to Neutrino Events in a Liquid Argon Time Projection Chamber,” *1611.05531* [physics.ins-det].
- [3] R. Acciarri *et al.* [SBN Collaboration], “A Proposal for a Three Detector Short-Baseline Neutrino Oscillation Program in the Fermilab Booster Neutrino Beam,” *arXiv:1503.01520* [physics.ins-det].
- [4] R. Acciarri *et al.* [DUNE Collaboration], “Long-Baseline Neutrino Facility (LBNF) and Deep Underground Neutrino Experiment (DUNE) : Volume 1: The LBNF and DUNE Projects,” *arXiv:1601.05471* [physics.ins-det].
- [5] E. Racah *et al.* “Revealing Fundamental Physics from the Daya Bay Neutrino Experiment using Deep Neural Networks,” *arXiv:1601.07621*.
- [6] A. Aurisano *et al.*, “A Convolutional Neural Network Neutrino Event Classifier,” 2016 *JINST* **11** P09001.
- [7] J. Renner *et al.* [NEXT Collaboration], “Background rejection in NEXT using deep neural networks,” *arxiv:1609.06202*.
- [8] LArSoft release *redmine* **05.08.00**
<https://cdcv.s.fnal.gov/redmine/projects/larsoft/wiki/ReleaseNotes050800>.
- [9] uboonecode release *redmine* **05.08.00**
<https://cdcv.s.fnal.gov/redmine/projects/uboonecode/wiki/ReleaseNotes050800>.
- [10] Genty, Terao and Wongjirad, LArCV: LArTPC Image Processing/Analysis Framework for Deep Learning <https://www.github.com/LArbys/LArCV>
- [11] Y. Jia *et al.*, “Caffe: Convolutional Architecture for Fast Feature Embedding,” *arXiv:1408.5093*, 2014.
- [12] Christian Szegedy *et al.*, “Going Deeper with Convolutions,” *arXiv:1409.4842*, 2014.
- [13] Ren, Shaoqing and He, Kaiming and Girshick, Ross and Sun, Jian, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks,” *NIPS* **28** pages 91-99, 2015.
- [14] Szegedy *et al.*, “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning,” *arXiv:1602.07261*.
- [15] Kaiming He *et al.*, “Deep Residual Learning for Image Recognition,” *arXiv:1512.03385*.
- [16] Rene Brun and Fons Rademakers, “ROOT – An object oriented data analysis framework,” *NIM* **A389** 81-86, 2015.
- [17] NVIDIA GeForce Titan X GPU Card Specification, <http://www.geforce.com/hardware/desktop-gpus/geforce-gtx-titan-x/specifications>
- [18] S. Agostinelli *et al.*, “GEANT4: A Simulation toolkit,” *NIM* **A506** 250-303, 2003.
- [19] MicroBooNE Collaboration, “TPC Noise Filtering,” MicroBooNE Public Note 1016-PUB.
<http://www-microboone.fnal.gov/publications/publicnotes/index.html>
- [20] MicroBooNE Collaboration. “TPC Signal Processing,” MicroBooNE Public Note 1017-PUB.
<http://www-microboone.fnal.gov/publications/publicnotes/index.html>
- [21] Montavon *et al.*, “Neural Networks: Tricks of the Trade,” Springer 2012.
- [22] A. A. Aguilar-Arevalo *et al.*, “The Neutrino Flux Prediction at MiniBooNE,” *PRD* **78**, 072002, 2009

- [23] C. Andreopoulos *et al.*, “The GENIE Neutrino Monte Carlo Generator,” *Int J Mod Phys A* **25** 87-104, 2010.
- [24] Ioffe *et al.*, “Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift,” *Proc. of the 32nd Intl. Conf. on Machine Learning* 37, 2015.
- [25] Srivastava *et al.*, “Dropout: A Simple Way to Prevent Neural Networks from Overfitting,” *Journal of Machine Learning Research* 15 1929-1958, 2014.
- [26] D. Silver *et al.*, “Mastering the game of Go with deep neural networks and tree search,” *Nature* 529, 484-489, 2016.
- [27] Tieleman, Tijmen and Hinton, Geoffrey. “Lecture 6.5-rmsprop: Divide the gradient by a running average of its recent magnitude,” *COURSERA: Neural Networks for Machine Learning*, 4, 2012.
- [28] MicroBooNE Collaboration, “Selection and kinematic properties of ν_μ charged-current inclusive events in 5e19 POT of MicroBooNE data,” <http://www-microboone.fnal.gov/publications/publicnotes/MICROBOONE-NOTE-1010-PUB.pdf>
- [29] Zheng *et al.*, “Improving the Robustness of Deep Neural Networks via Stability Training,” *arXiv:1604.04326*.

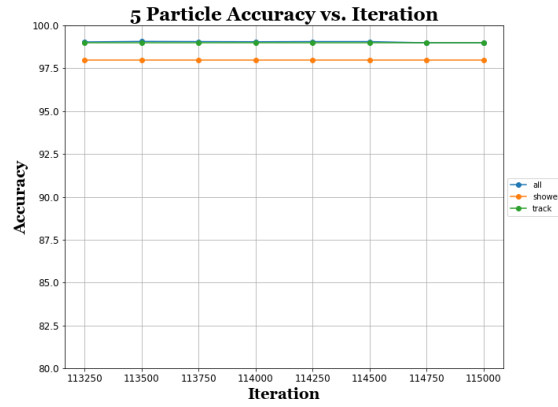
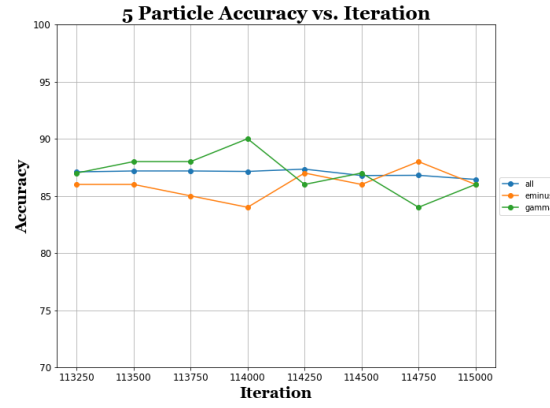
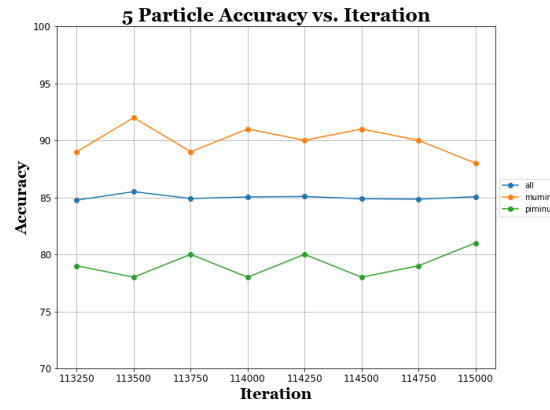
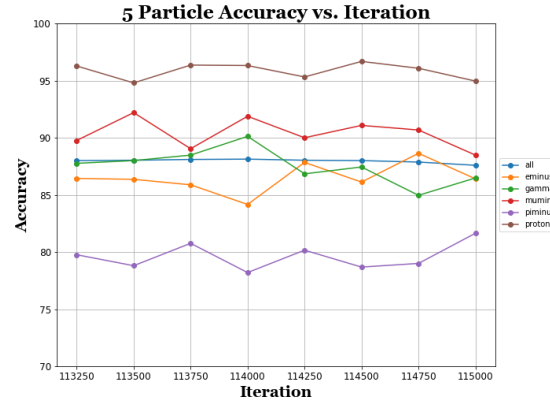


Figure 2: This is the five particle classification, and μ^- , π^- , e , γ , proton and shower-track separations for the vgg16b network.