Appendix: Material phase prediction for Li-ion Battery Reconstruction using Hierarchical Curriculum Learning

[1] Anika Tabassum, [2] Nikhil Muralidhar, [1] Ramakrishnan Kannan, [1] Srikanth Allu
[1] *Oak Ridge National Laboratory*
[2] *Department of Computer Science, Stevens Institute of Technology*
Email: {tabassuma, allus , kannanr}@ornl.gov, nmurali1@stevens.edu

I. Extension to Sec. IV

TABLE I: **LCGAR Architecture.**

Layer Type	Specifications
Convolution layer	Kernel $(3 \times 3,)$, 32 channels, padding 1
Convolution layer	Kernel $(3 \times 3,)$, 32 channels, padding 1
Maxpool	Kernel (2×2)
Convolution layer	Kernel $(3 \times 3,)$, 16 channels, padding 1
Convolution layer	Kernel $(3 \times 3,)$, 16 channels, padding 1
Maxpool	Kernel (2×2)
Convolution layer	Kernel $(1 \times 1,)$, 8 channels
Convolution layer	Kernel $(1 \times 1,)$, 8 channels
Concat: Flatten + UGSM logits $\mathbf{F}^j$	
Fully Connected Layer + dropout	Size 128, dropout (20%)
Batch-normalization	Size 128
Softmax	Number of class (3)

Table I shows the overview of LCGAR architecture.

II. Extension to Sec. V

Model Training: Both UGSM and LCGAR are trained with weighted cross-entropy loss due to heavy class imbalance in the ground-truth (on average C: 26%, Ni: 60%, Pore: 14%). In all the experiments, for LCGAR, we normalize input image between [-1, 1] and use activation *LeakyReLU*, which yields the best results. Our UGSM model is trained on 100 epochs with batch size 20, while HCL-IDK and LCGAR were trained for 15 epochs with batch size 1024.

Data Collection Process: X-Ray Computed Tomography (XCT) and similar techniques do not provide information on carbon binder distribution as it is "transparent" for the X-Ray. It thus becomes indistinguishable from the pores in the electrodes. For collecting the ground-truth labels, to identify all the material constituents, we utilized cross-sectioning with Focused-Ion beam (FIB/SEM) experiments for imaging and Energy Dispersive X-Ray Spectroscopy (EDS) for chemical mapping of electrode cross-sections [1]. The slicing was performed using Hitachi NB500 dual-beam FIB/SEM. The cross-sections were at a distance of every $200nm$ resulting in a total depth of $26\mu m$ resulting in 133 images.

Pre-processing: In this paper, we construct our data-set $\mathcal{D}$ using the corpus of 133 images collected through cross-sections of an electrode. Each low-contrast image $\mathbf{X}$ in the corpus consists of 224×256 pixels. The corresponding ground-truth (GT) $\mathbf{Y}$ contains 224×256 material constituents consisting of pores, carbon binder (C), and Nickel (Ni). We leverage a data augmentation technique for efficient training with few data samples. Our technique is: First, for every image $\mathbf{X}$ and the corresponding GT, we obtain k different images, removing a row of pixels from the top. Next, we resize each k images into original size, i.e., 224×256. We choose $k = 10$, as this is sufficient for training data. Finally, to smooth the resized GT for the corresponding k images, we use existing knowledge provided by domain experts: A pixel with Carbon (C) can not exist surrounding Ni pixels and vice-versa for a pixel with Ni. Using k nearest-neighbor rule, we remove noise from GT, i.e., if most surrounding pixels (nearest neighbors) around a C pixel are Ni, the GT label is changed to Ni.

Measure of Success:

- F1-score: Our goal is to measure the overall prediction for each class c, i.e., pore, carbon (C), and nickel (Ni) for the unseen datasets.
- *Pixel accuracy (ACC)*: Fraction of the number of pixels that are predicted correctly among total pixels (in %).To measure smooth predictions [2], We aim to evaluate pixel accuracy for k best performing and worst performing predictions to evaluate smoothness of material phase predictions.
- *Mean intersection over union (mIU)*: We intend to quantify predictions from standard practice of image segmentation models [3], [4]. Suppose, t_c be the total number of pixels labeled as class c, n_{jc} be the number of pixels of class c predicted as class j, and $|C|$ are the total number of classes. mIU $= \frac{1}{|C|} \sum_c \frac{n_{cc}}{t_c + \sum_{j \neq c} n_{jc}}$
- *Frequency weighted intersection over union (fIU)*: To quantify predictions in presence of class imbalance, we incorporate *fIU* from the standard practice for image segmentation models [3]. If t be the total number of pixels, fIU $= \frac{1}{t} \sum_c \frac{t_c * n_{cc}}{t_c + \sum_{j \neq c} n_{jc}}$

Table II shows MatPhase performance compared with baselines.

References

[1] S. Vierrath *et al.*, "Morphology of nanoporous carbon-binder domains in li-ion batteries—a fib-sem study," *Electrochemistry Communications*,

TABLE II: Performance of MatPhase comparing with the baselines in terms of F1-score per class, accuracy (%) on top 5 best performing, and lowest performing test sets, *mIU*, and *fIU*. For all metrics, high score yields the better result. MatPhase outperforms all the models (best values in bold).

Model	F1			ACC(%)	ACC(%)	mIU	fIU
	Pore	C	Ni	Best 5	Lowest 5		
DeepLabV3 [5]	0.85	**0.86**	0.87	90± 3.85	89± 1.87	0.76	0.83
MANet [6]	0.81	0.79	0.93	90.7± 3.15	85.6± .0002	0.74	0.82
FCN [3]	0.66	0.17	0.88	80± 0.2	75.8± 1.2	0.45	0.64
SegNet [4]	0.81	0.78	0.94	91.7± 0.1	85.6± 1.6	0.74	0.83
U-Net [7]	0.85	0.82	0.94	91.2± .09	90.03± 1	0.77	0.84
U-Net++ [8]	0.76	0.53	0.78	72.4± 1.46	67.8± 3.09	0.54	0.59
MCD-U-Net [9]	**0.86**	0.82	0.94	91.5± 0.09	90.5± 0.6	0.78	0.84
Local-U-Net	**0.86**	0.82	0.94	91.5± 0.04	90.5± 0.5	0.78	0.84
ResNet-18 [10]	**0.86**	0.83	**0.95**	92.7 ± .03	90.6 ± 0.02	0.79	0.86
Adapted-LCGAR	0.85	0.84	**0.95**	93 ± 0.02	91.8 ± 1.2	0.79	0.86
MatPhase	**0.86**	0.85	**0.95**	**93.2** ± 0.02	**91.9** ± 1.3	**0.80**	**0.87**

vol. 60, pp. 176–179, 2015.

[2] S. Sinha *et al.*, "Curriculum by smoothing," *Proc. of NeurIPS*, 2020.

[3] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proceedings of the IEEE CVPR*, 2015, pp. 3431–3440.

[4] V. Badrinarayanan *et al.*, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation," *IEEE transactions on pattern analysis and machine intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.

[5] L.-C. Florian and S. H. Adam, "Rethinking atrous convolution for semantic image segmentation," in *Conference on CVPR. IEEE/CVF*, vol. 6, 2017.

[6] T. Fan *et al.*, "Ma-net: A multi-scale attention network for liver and tumor segmentation," *IEEE Access*, vol. 8, pp. 179 656–179 665, 2020.

[7] O. Ronneberger *et al.*, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention*, ser. LNCS, vol. 9351. Springer, 2015, pp. 234–241.

[8] Z. Zhou *et al.*, "Unet++: A nested u-net architecture for medical image segmentation," in *Deep learning in medical image analysis and multimodal learning for clinical decision support*. Springer, 2018, pp. 3–11.

[9] T. DeVries and G. W. Taylor, "Leveraging uncertainty estimates for predicting segmentation quality," *arXiv preprint arXiv:1807.00502*, 2018.

[10] K. He *et al.*, "Deep residual learning for image recognition," in *Proceedings of the IEEE CVPR*, 2016, pp. 770–778.