Supplemental Figures

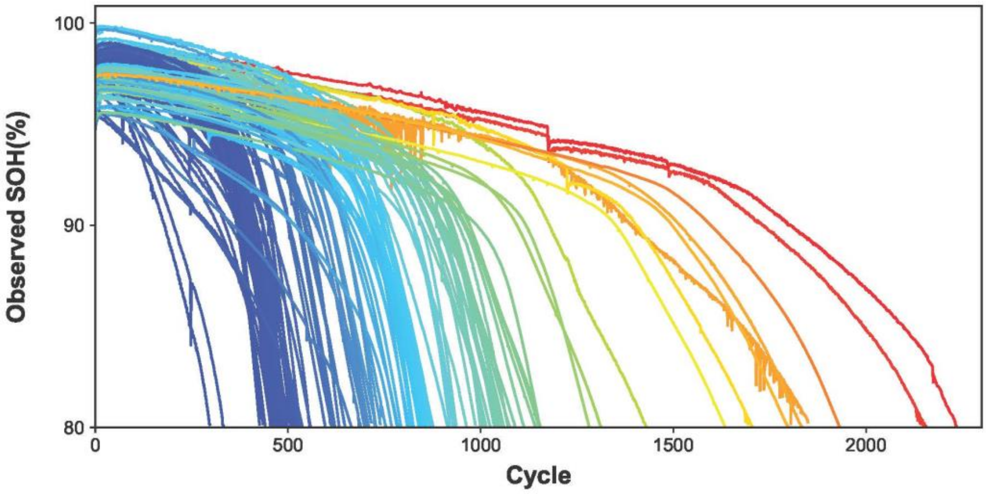

Figure S1. Observed SOH of all the LFP/ graphite batteries in the dataset [1]. The color is scaled by cycle life for each battery.

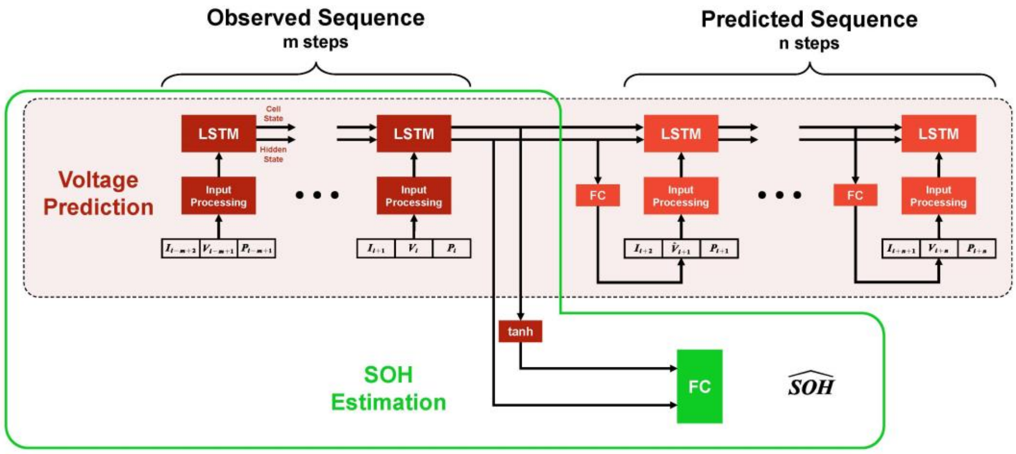

Figure S2. Transfer learning framework. The voltage prediction task was initially trained. Subsequently the SOH prediction task is trained under the constraint that the shared part, namely the observed sequence of voltage prediction, was frozen. The input of SOH estimation then becomes:

$$\xi_{SOH} = [h_{i,I},\ \tanh(C_{i,I})]^{\mathrm{T}}$$

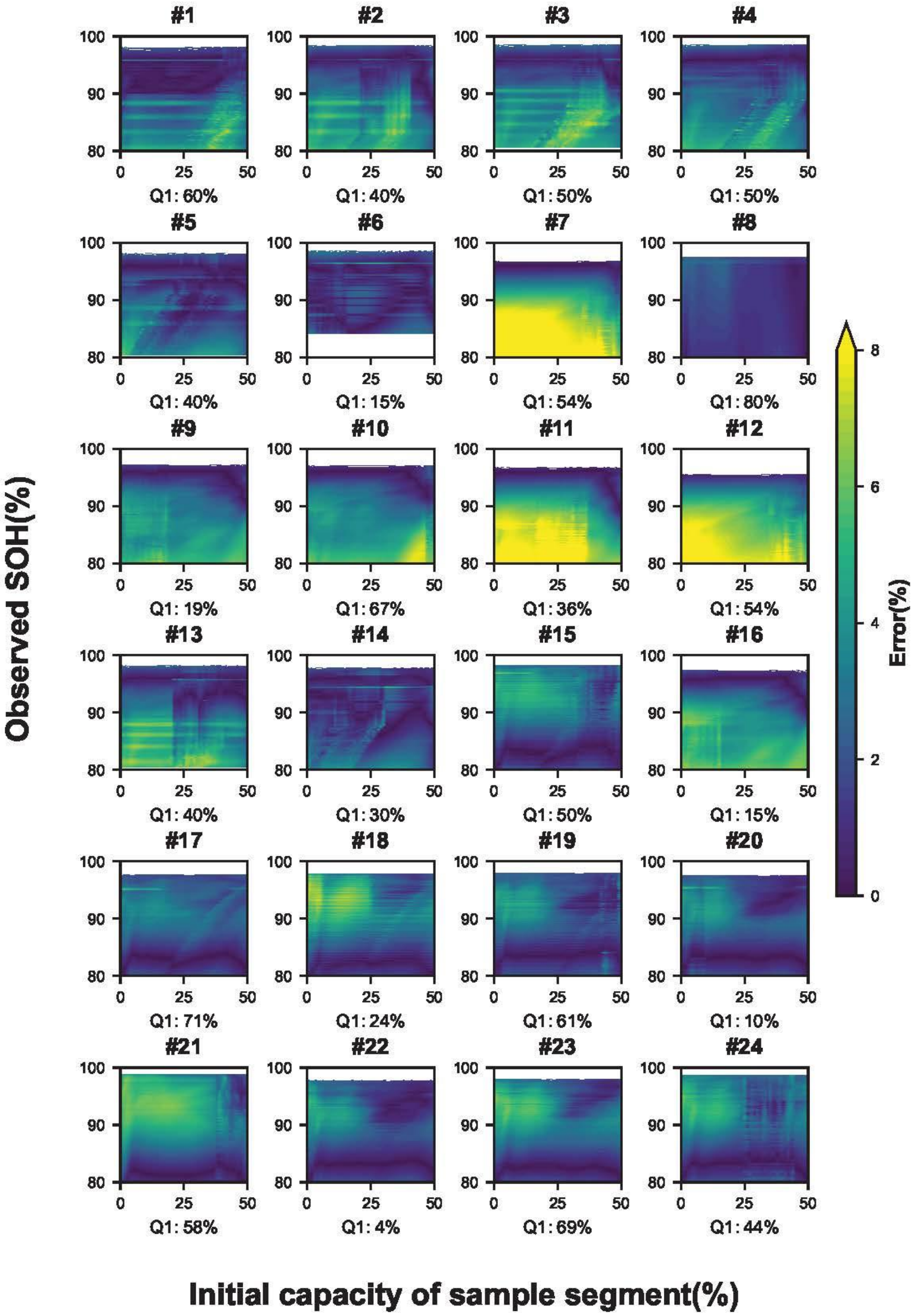

Figure S3. The SOH estimation errors of CNN using the PyTorch library [2] resulted in an MAE of 1.91% and an RMSE of 2.47%. Both were highest among all comparison methods. CNN is more concerned with the relationship between the shape of the input and the result, which is not the fundamental logic of battery.

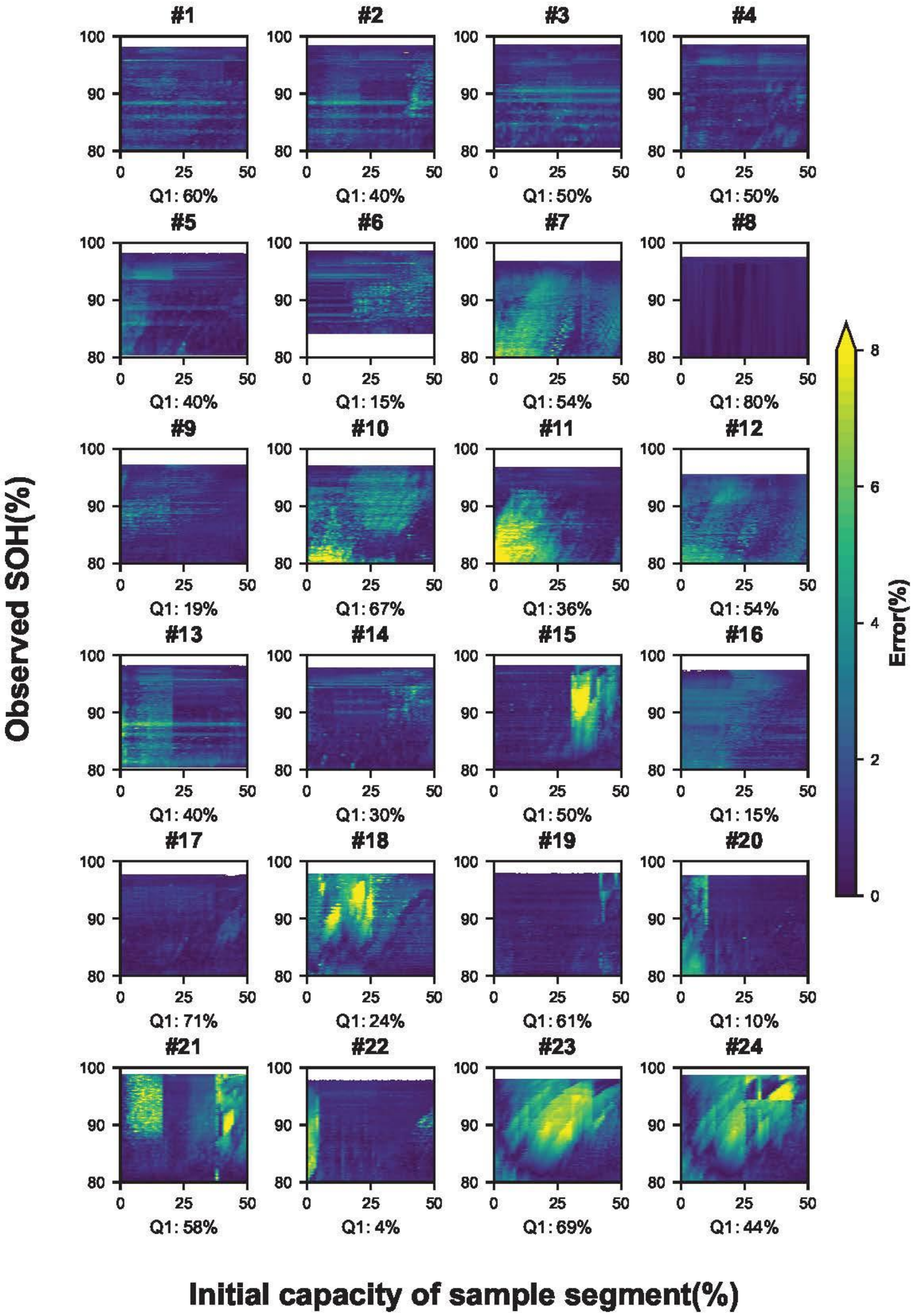

Figure S4. The SOH estimation errors of RF using the scikit-learn library [3] resulted in an MAE of 1.04% and an RMSE of 1.57%. It is can be spotted that the errors were notably higher when the sample segments contain the current switch position, especially for those batteries charged by a strategy specified for themselves. This showed that RF was not competent to learn the underlying mechanism and estimate SOH under uncertain charging strategies. RF is able to identify logical relationships to some extent, but does not analyze sequences from a temporal perspective.

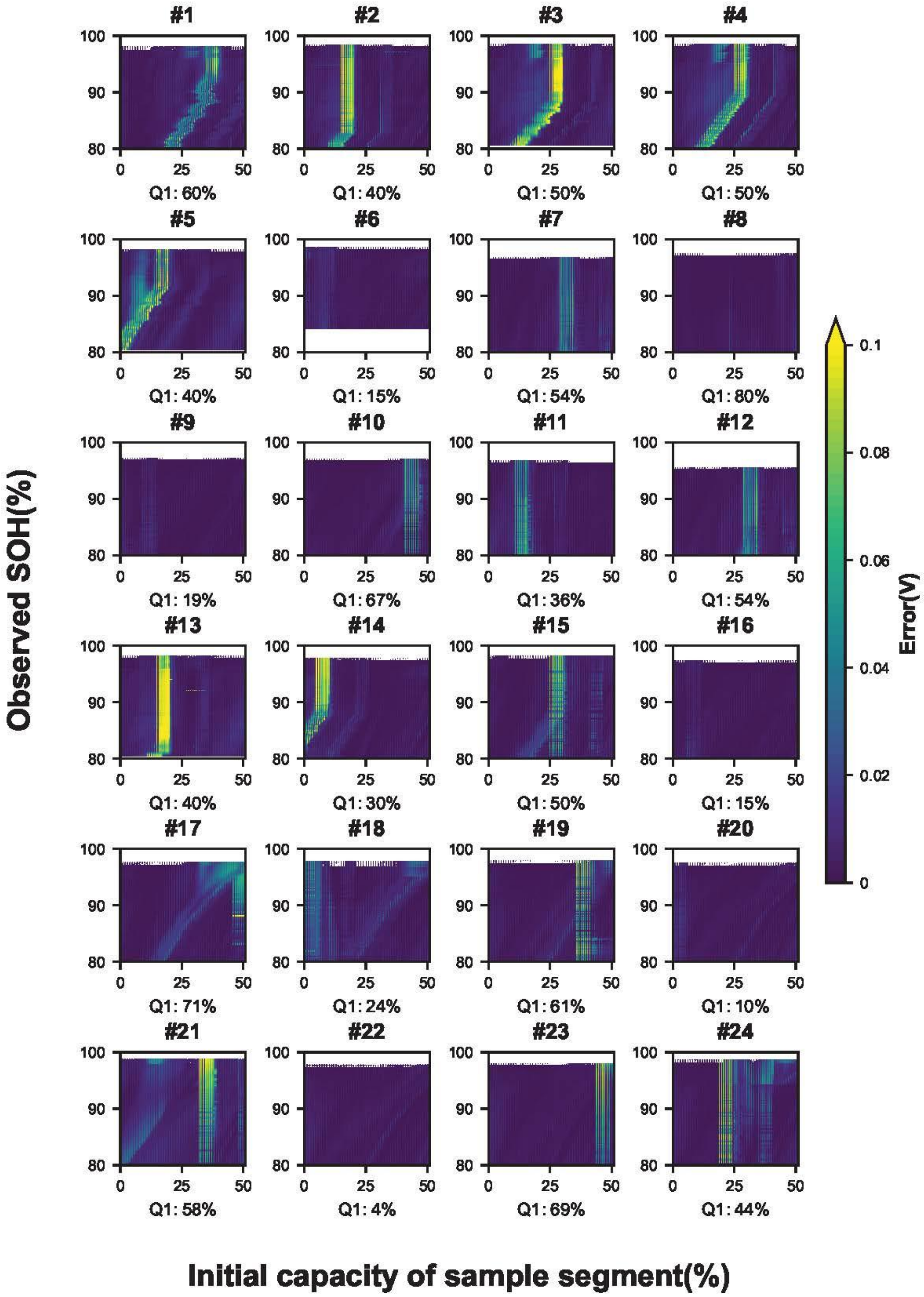

Figure S5. The voltage prediction errors of transfer learning using the PyTorch library resulted in an MAE of 0.0041V and an RMSE of 0.0109V. The voltages were predicted directly and the results showed similar characteristics with MTL. The errors at the switch positions of the charging currents were higher and formed lines due to the interpolation of data.

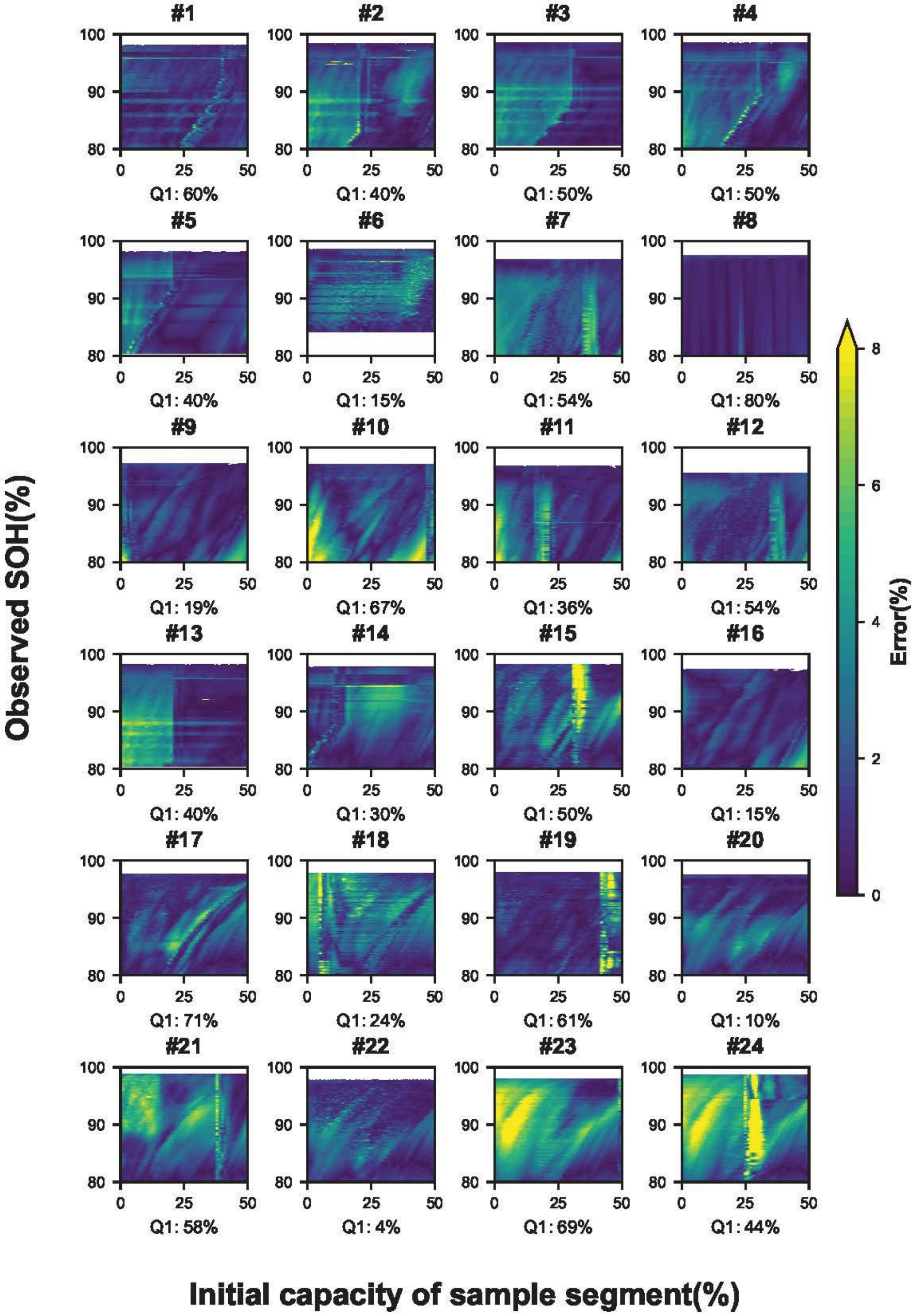

Figure S6. The results of SOH estimation using transfer learning, with the MAE of 1.39% and the RMSE of 1.81%. It is similar to MTL that the estimation errors right after the switch position are high, but then decrease rapidly. It was a proof of preliminary capacity to recognize the polarization state and eliminate its influence while estimating SOH.

Supplemental Tables

Table S1. The charging strategies for training set including 99 batteries. C1 and C2 represented the charging rates for the first and second steps, respectively. Q1 represented the nominal capacity percentage at which the charging rate was switched. The second step was terminated when charged 80% of the nominal capacity.

Training index	C1	Q1	C2
#1	'3.6C'	'80%'	'3.6C'
#2	'3.6C'	'80%'	'3.6C'
#3	'3.6C'	'80%'	'3.6C'
#4	'4C'	'80%'	'4C'
#5	'4C'	'80%'	'4C'
#6	'4.4C'	'80%'	'4.4C'
#7	'4.8C'	'80%'	'4.8C'
#8	'4.8C'	'80%'	'4.8C'
#9	'5.4C'	'40%'	'3.6C'
#10	'5.4C'	'50%'	'3C'
#11	'5.4C'	'60%'	'3C'
#12	'5.4C'	'60%'	'3C'
#13	'5.4C'	'60%'	'3.6C'
#14	'5.4C'	'70%'	'3C'
#15	'5.4C'	'70%'	'3C'
#16	'5.4C'	'80%'	'5.4C'
#17	'5.4C'	'80%'	'5.4C'
#18	'6C'	'30%'	'3.6C'
#19	'6C'	'40%'	'3C'
#20	'6C'	'40%'	'3.6C'
#21	'6C'	'50%'	'3C'
#22	'6C'	'50%'	'3.6C'
#23	'6C'	'60%'	'3C'
#24	'6C'	'60%'	'3C'
#25	'7C'	'30%'	'3.6C'
#26	'7C'	'40%'	'3C'
#27	'7C'	'40%'	'3.6C'
#28	'7C'	'40%'	'3.6C'
#29	'8C'	'15%'	'3.6C'
#30	'8C'	'25%'	'3.6C'
#31	'8C'	'25%'	'3.6C'
#32	'8C'	'35%'	'3.6C'
#33	'8C'	'35%'	'3.6C'
#34	'1C'	'4%'	'6C'
#35	'2C'	'2%'	'5C'
#36	'2C'	'7%'	'5.5C'

#37	'3.6C'	'22%'	'5.5C'
#38	'3.6C'	'2%'	'4.85C'
#39	'3.6C'	'30%'	'6C'
#40	'3.6C'	'9%'	'5C'
#41	'4C'	'13%'	'5C'
#42	'4C'	'31%'	'5C'
#43	'4C'	'40%'	'6C'
#44	'4C'	'4%'	'4.85C'
#45	'4.4C'	'47%'	'5.5C'
#46	'4.4C'	'55%'	'6C'
#47	'4.4C'	'8%'	'4.85C'
#48	'4.65C'	'19%'	'4.85C'
#49	'4.8C'	'80%'	'4.8C'
#50	'4.8C'	'80%'	'4.8C'
#51	'4.9C'	'27%'	'4.75C'
#52	'4.9C'	'69%'	'4.25C'
#53	'5.2C'	'37%'	'4.5C'
#54	'5.2C'	'58%'	'4C'
#55	'5.2C'	'66%'	'3.5C'
#56	'5.6C'	'25%'	'4.5C'
#57	'5.6C'	'38%'	'4.25C'
#58	'5.6C'	'47%'	'4C'
#59	'5.6C'	'5%'	'4.75C'
#60	'5.6C'	'65%'	'3C'
#61	'6C'	'20%'	'4.5C'
#62	'6C'	'31%'	'4.25C'
#63	'6C'	'40%'	'4C'
#64	'6C'	'52%'	'3.5C'
#65	'6C'	'60%'	'3C'
#66	'5C'	'67%'	'4C'
#67	'5.3C'	'54%'	'4C'
#68	'5.6C'	'36%'	'4.3C'
#69	'5.6C'	'19%'	'4.6C'
#70	'5.6C'	'36%'	'4.3C'
#71	'3.7C'	'31%'	'5.9C'
#72	'4.8C'	'80%'	'4.8C'
#73	'5C'	'67%'	'4C'
#74	'5.3C'	'54%'	'4C'
#75	'4.8C'	'80%'	'4.8C'
#76	'5.6C'	'19%'	'4.6C'
#77	'5.6C'	'36%'	'4.3C'
#78	'5.6C'	'36%'	'4.3C'
#79	'5.9C'	'15%'	'4.6C'

#80	'4.8C'	'80%'	'4.8C'
#81	'5.3C'	'54%'	'4C'
#82	'5.6C'	'19%'	'4.6C'
#83	'5.6C'	'36%'	'4.3C'
#84	'3.7C'	'31%'	'5.9C'
#85	'5.9C'	'60%'	'3.1C'
#86	'5C'	'67%'	'4C'
#87	'5.3C'	'54%'	'4C'
#88	'5.6C'	'19%'	'4.6C'
#89	'3.7C'	'31%'	'5.9C'
#90	'5.3C'	'54%'	'4C'
#91	'5.9C'	'60%'	'3.1C'
#92	'5C'	'67%'	'4C'
#93	'5.6C'	'19%'	'4.6C'
#94	'5.6C'	'36%'	'4.3C'
#95	'5C'	'67%'	'4C'
#96	'5.6C'	'19%'	'4.6C'
#97	'5.6C'	'36%'	'4.3C'
#98	'5.3C'	'54%'	'4C'
#99	'4.8C'	'80%'	'4.8C'

Table S2a. Hyperparameters of CNN

1st convolutional layer kernel size	3
1st convolutional layer stride	3
1st convolutional layer padding	1
1st pooling layer kernel size	3
1st pooling layer stride	3
2nd convolutional layer kernel size	3
2nd convolutional layer stride	3
2nd convolutional layer padding	1
2nd pooling layer kernel size	3
2nd pooling layer stride	3
1st FC layer weight matrix size	128*1536
2ne FC layer weight matrix size	64*128
3rd FC layer weight matrix size	1*64
Training epochs	100
Batch size	10000

Table S2b. Hyperparameters of RF

The number of trees in the forest	10
The minimum number of samples to split an internal node	2
Training epoch	100

Table S2c. Hyperparameters of transfer learning

l	500
d	0.0004
m	100 (20% nominal capacity)
n	35 (7% nominal capacity)
Input process W_1 size	128*3
Input process W_2 size	32*128
Input process W_3 size	32*32
Input process W_4 size	32*32
LSTM Hidden states h_i dimension	64
LSTM Cell states C_i dimension	64
weight matrix size of 1st FC layer for SOH estimation	128*128
weight matrix size of 2nd FC layer for SOH estimation	128*128
weight matrix size of 3rd FC layer for SOH estimation	128*1
Training epochs for voltage prediction task	30
Training epochs for SOH estimation task	200
Batch size for voltage prediction task	20000
Batch size for SOH estimation task	10000

Supplemental References

[1] Severson, K.A., Attia, P.M., Jin, N., Perkins, N., Jiang, B., Yang, Z., Chen, M.H., Aykol, M., Herring, P.K., and Fraggedakis, D.J.N.E. (2019). Data-driven prediction of battery cycle life before capacity degradation. Nature Energy 4, 383-391.

[2] https://pytorch.org/

[3] https://scikit-learn.org/stable/