Supplementary information

Real-time personalized health status prediction of lithium-ion batteries using deep transfer learning

Guijun Ma,[a] Songpei Xu,[b] Benben Jiang,[c] Cheng Cheng,[b] Xin Yang,[d]

Yue Shen,[e] Tao Yang,[f] Yunhui Huang,[e] Han Ding[a] and Ye Yuan*[b]

[a] School of Mechanical Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China.

[b] School of Artificial Intelligence and Automation, Huazhong University of Science and Technology, Wuhan 430074, China. E-mail: yye@hust.edu.cn

[c] Department of Automation, Tsinghua University, Beijing 100084, China.

[d] College of Electrical and Information Engineering, Hunan University, Changsha, 410082, China.

[e] School of Materials Science and Engineering, Huazhong University of Science and Technology, Wuhan 430074, China.

[f] State Key Laboratory of Synthetical Automation for Process Industries, Northeastern University, Shenyang 110819, China.

Supplemental Figures

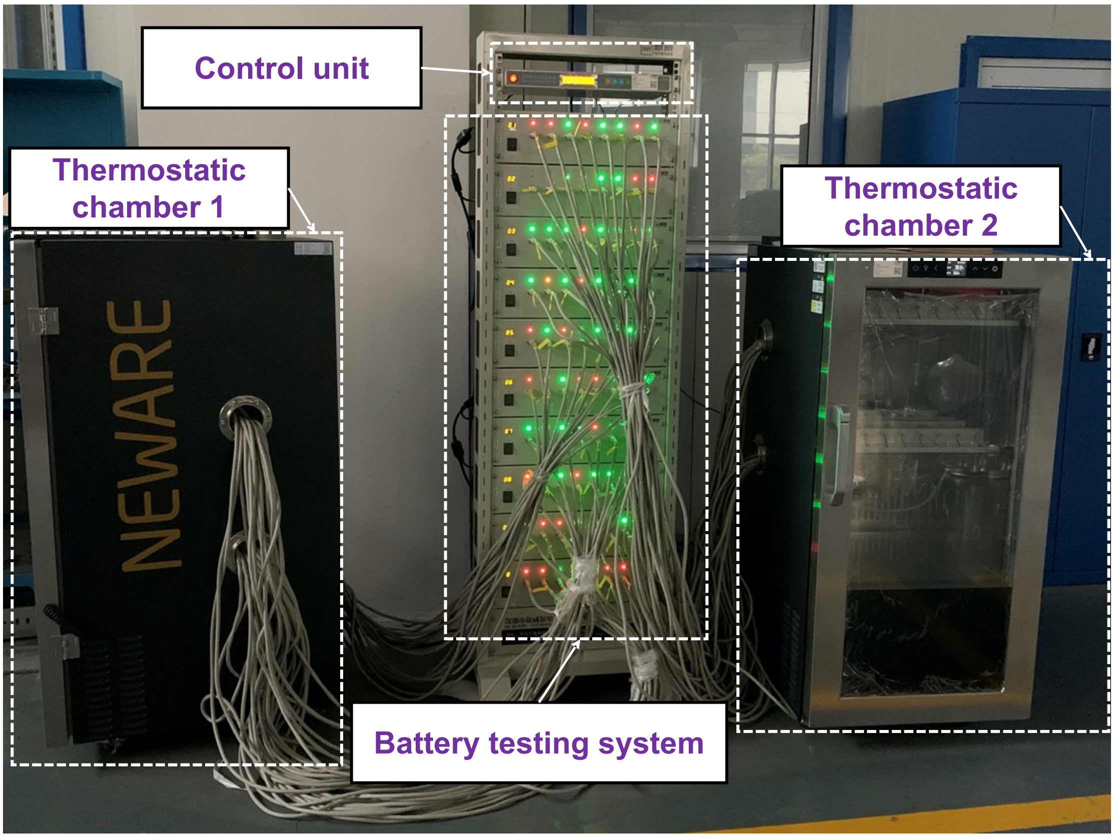

Fig. S1 The developed battery degradation platform in our laboratory. The platform consists of a control unit, a battery testing system (80 channels), two thermostatic chambers and 77 lithium-iron-phosphate/graphite A123 APR18650M1A cells.

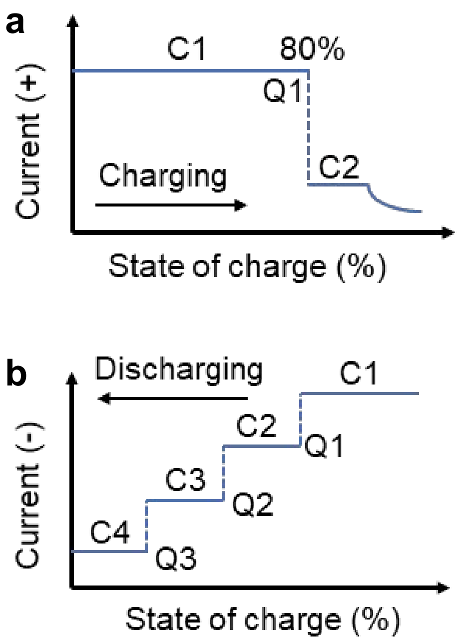

Fig. S2 The illustration of charge protocol and discharge protocol in the battery degradation experiment. (a) An identical charge protocol C1 (5C, 0% SOC-80% SOC)-C2 (1C, 80% SOC-the first 3.6 V) followed by a constant-voltage charge (3.6 V-C/20) for cells. (b) Four-step discharge from 100% SOC to a cutoff voltage of 2 V. Diverse discharge protocols C1 (100% SOC-60% SOC)-C2 (60% SOC-40% SOC)-C3 (40% SOC-20% SOC)-C4 (20% SOC-2 V) for different cells, where C1-C4 represent constant-current discharge rates at the four steps, respectively.

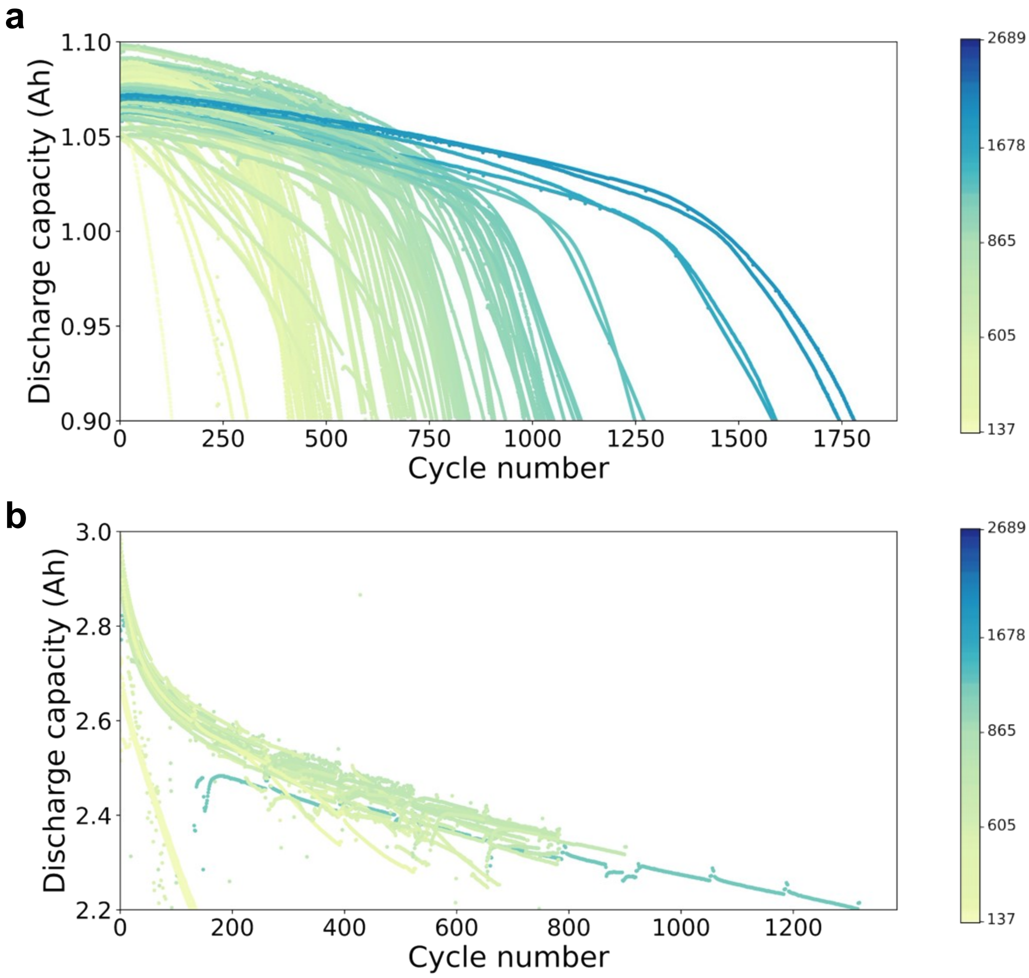

Fig. S3 Discharge capacity versus cycle number of (a) LFP/graphite cells developed by Severson et al.[1] and (b) 22 NMC/graphite cells developed by Preger et al.[2]

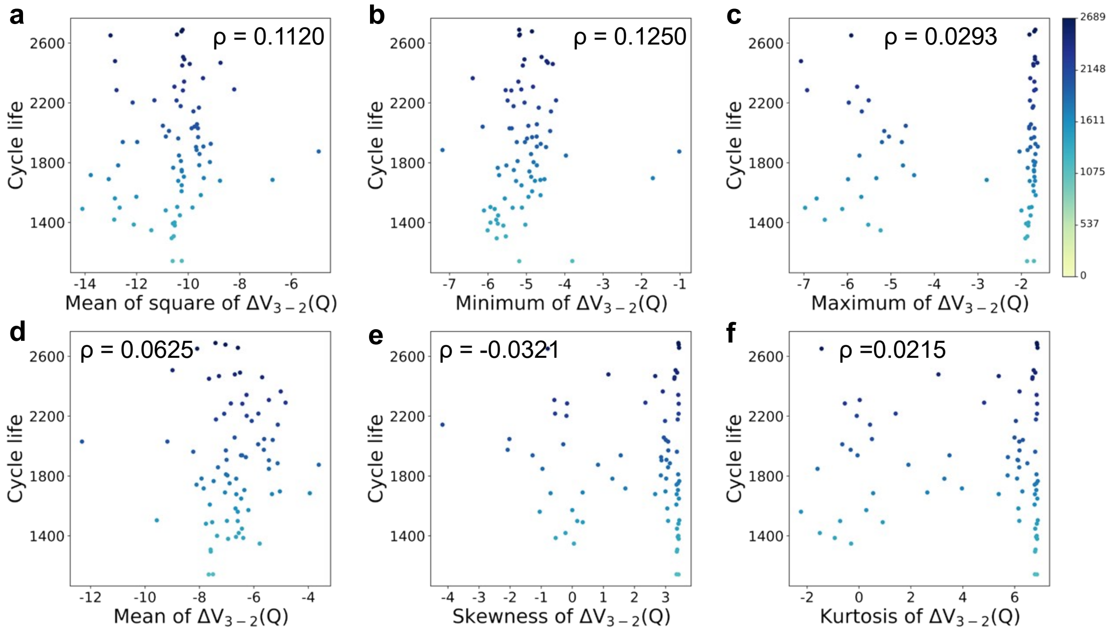

Fig. S4 The cycle life versus handcrafted features[3] which are extracted from the charge voltage difference between the 3rd charge-discharge cycle and the 2nd charge-discharge cycle along the battery degradation process. Cyle life versus (a) mean of square of $\Delta V_{3-2}(Q)$, (b) minimum of $\Delta V_{3-2}(Q)$, (c) maximum of $\Delta V_{3-2}(Q)$, (d) mean of $\Delta V_{3-2}(Q)$, (e) skewness of $\Delta V_{3-2}(Q)$ and (f) kurtosis of $\Delta V_{3-2}(Q)$. The correlation coefficient between each handcrafted feature and cycle life is marked in each subfigure.

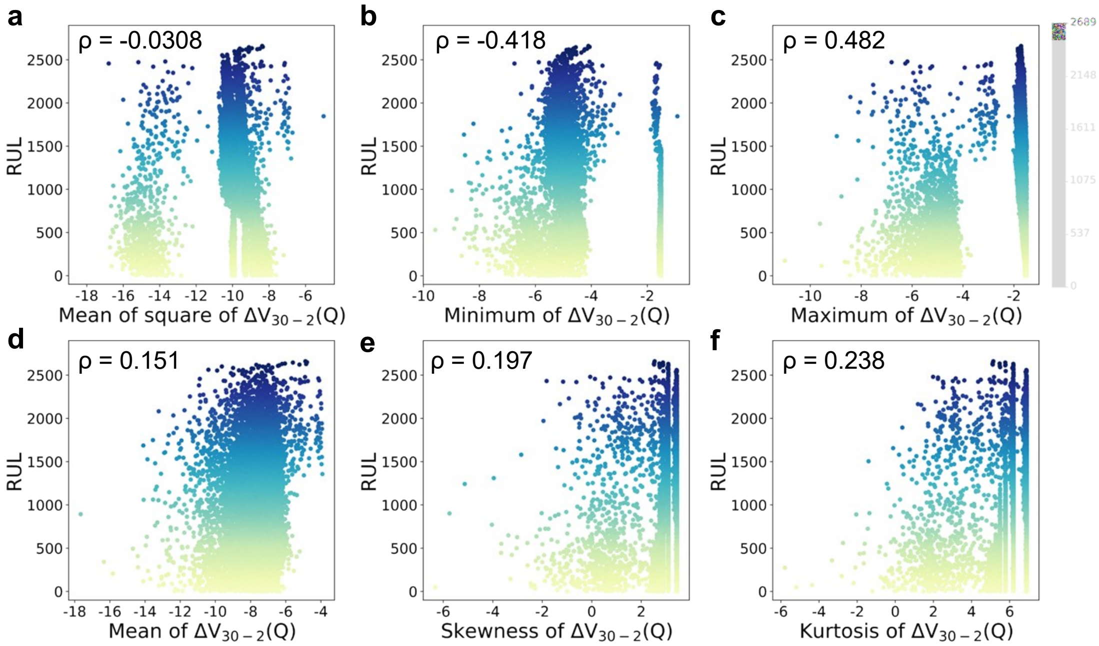

Fig. S5 The RUL versus handcrafted features which are extracted from the charge voltage difference between the 30th charge-discharge cycle and the 2nd charge-discharge cycle of every 30 cycles. RUL versus (a) mean of square of $\Delta V_{30-2}(Q)$, (b) minimum of $\Delta V_{30-2}(Q)$, (c) maximum of $\Delta V_{30-2}(Q)$, (d) mean of $\Delta V_{30-2}(Q)$, (e) skewness of $\Delta V_{30-2}(Q)$ and (f) kurtosis of $\Delta V_{30-2}(Q)$. The correlation coefficient between each handcrafted feature and cycle life is marked in each subfigure.

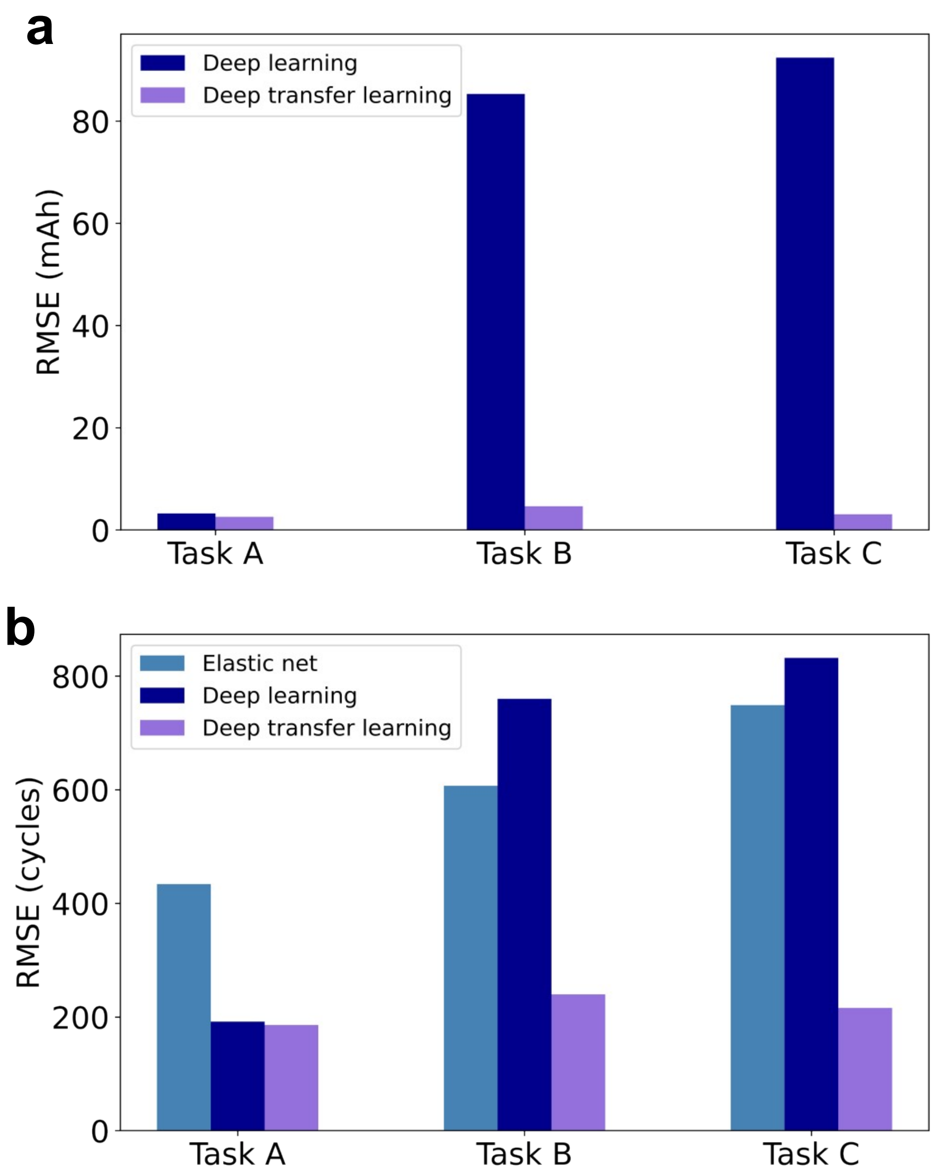

Fig. S6 The comparison of our deep transfer learning model with benchmark models. RMSE is used as the comparison metric for (a) capacity estimation and (b) RUL prediction. Comparisons for three different tasks are performed, respectively.

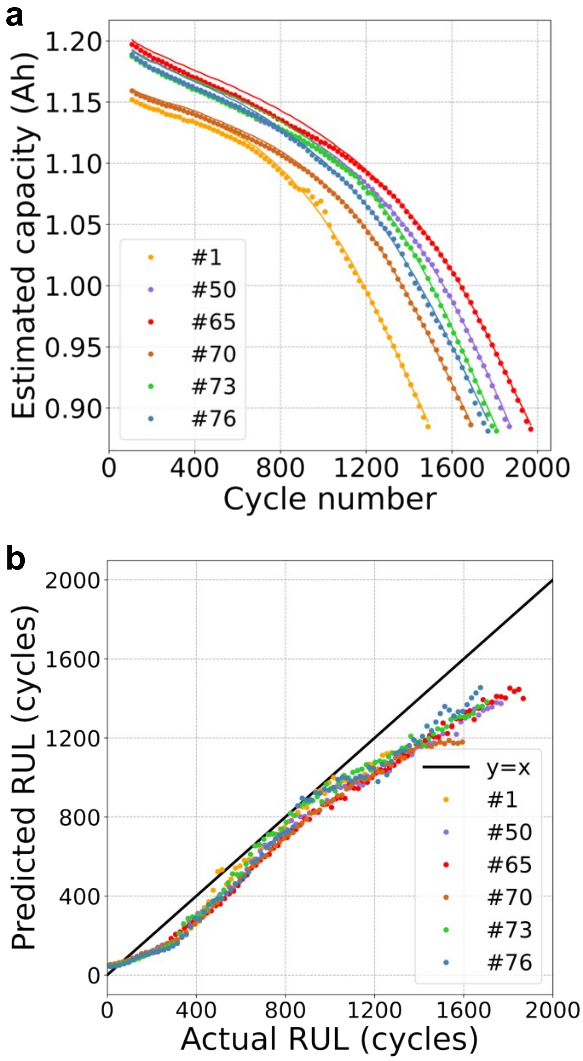

Fig. S7 Predicted health status versus actual health status for 6 testing protocols (protocols #1, #50, #65, #70, #73 and #76). The model is pre-trained by a public dataset with LFP/graphite cells that has various charge protocols, but an identical discharge protocol. (a) The predicted capacity versus actual capacity of the 6 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 6 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 6 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

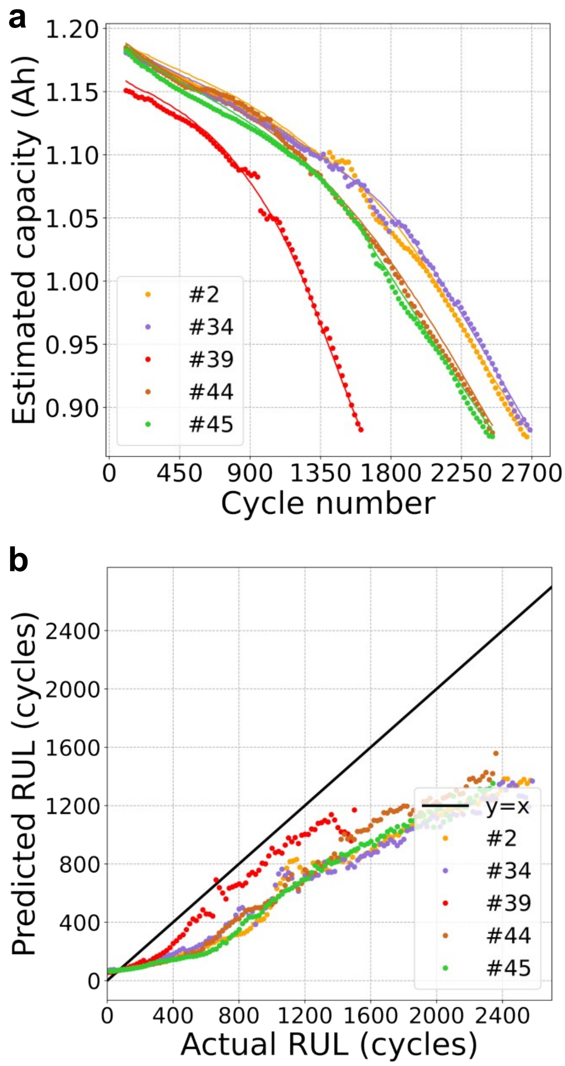

Fig. S8 Predicted health status versus actual health status for 5 testing protocols (protocols #2, #34, #39, #44 and #45). The model is pre-trained by a public dataset with LFP/graphite cells that has various charge protocols, but an identical discharge protocol. (a) The predicted capacity versus actual capacity of the 5 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 5 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 5 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

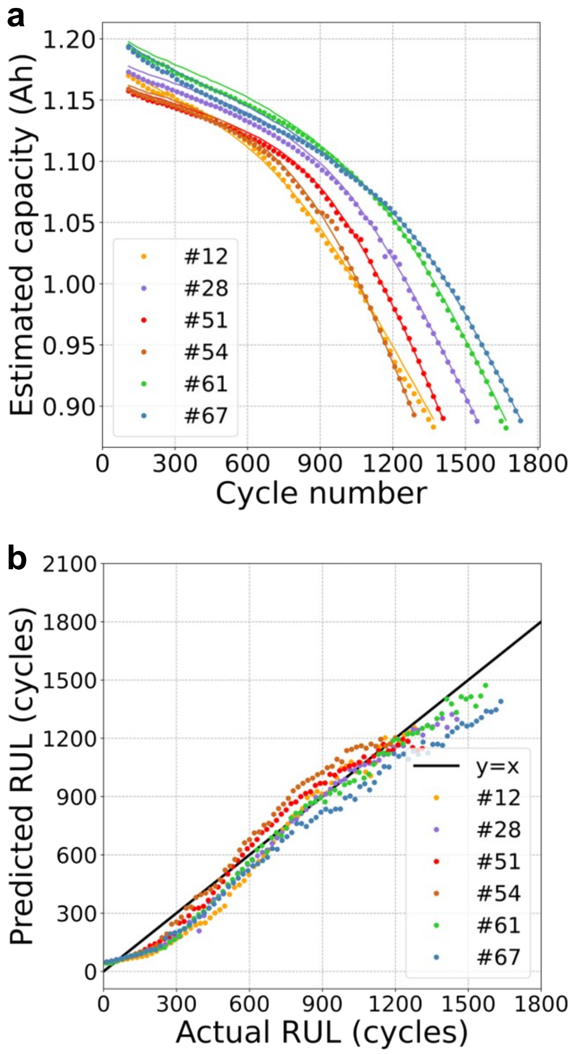

Fig. S9 Predicted health status versus actual health status for 6 testing protocols (protocols #12, #28, #51, #54, #61 and #67). The model is pre-trained by a public dataset with LFP/graphite cells that has various charge protocols, but an identical discharge protocol. (a) The predicted capacity versus actual capacity of the 6 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 6 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 6 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

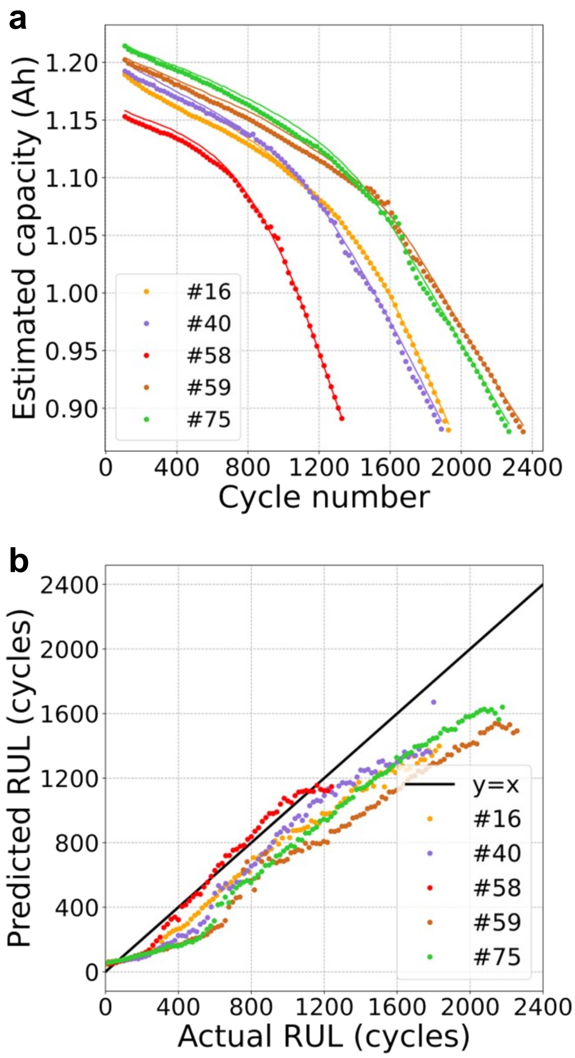

Fig. S10 Predicted health status versus actual health status for 5 testing protocols (protocols #16, #40, #58, #59 and #75). The model is pre-trained by a public dataset with LFP/graphite cells that has various charge protocols, but an identical discharge protocol. (a) The predicted capacity versus actual capacity of the 5 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 5 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 5 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

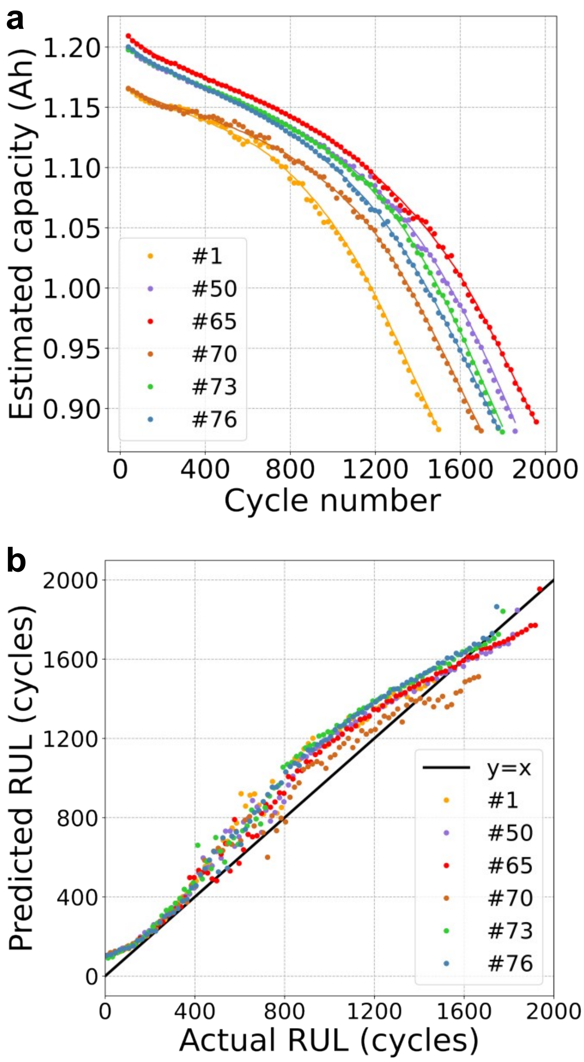

Fig. S11 Predicted health status versus actual health status for 6 testing protocols (protocols #1, #50, #65, #70, #73 and #76). The model is pre-trained by a public dataset with 22 NMC/graphite cells. (a) The predicted capacity versus actual capacity of the 6 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 6 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 6 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

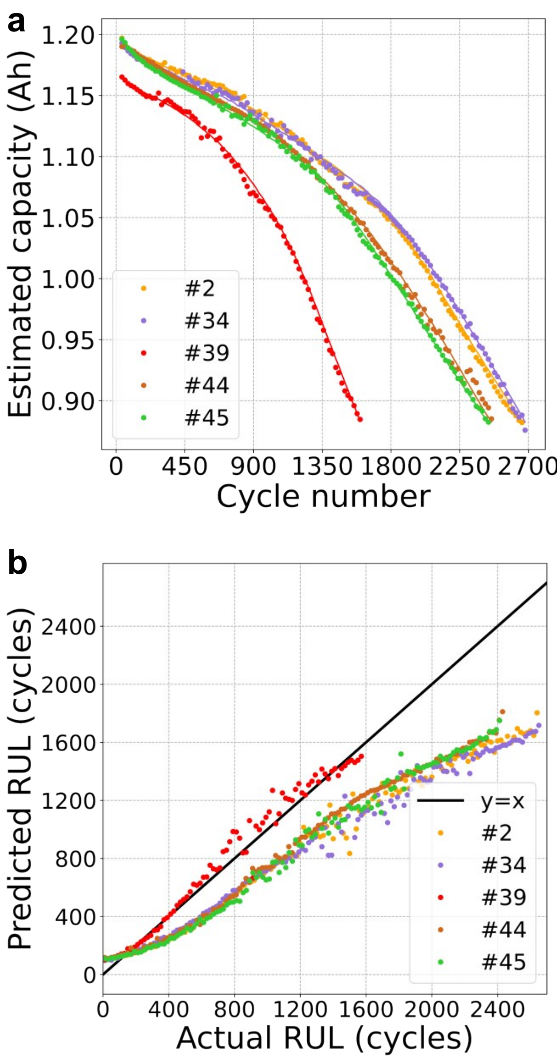

Fig. S12 Predicted health status versus actual health status for 5 testing protocols (protocols #2, #34, #39, #44 and #45). The model is pre-trained by a public dataset with 22 NMC/graphite cells. (a) The predicted capacity versus actual capacity of the 5 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 5 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 5 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

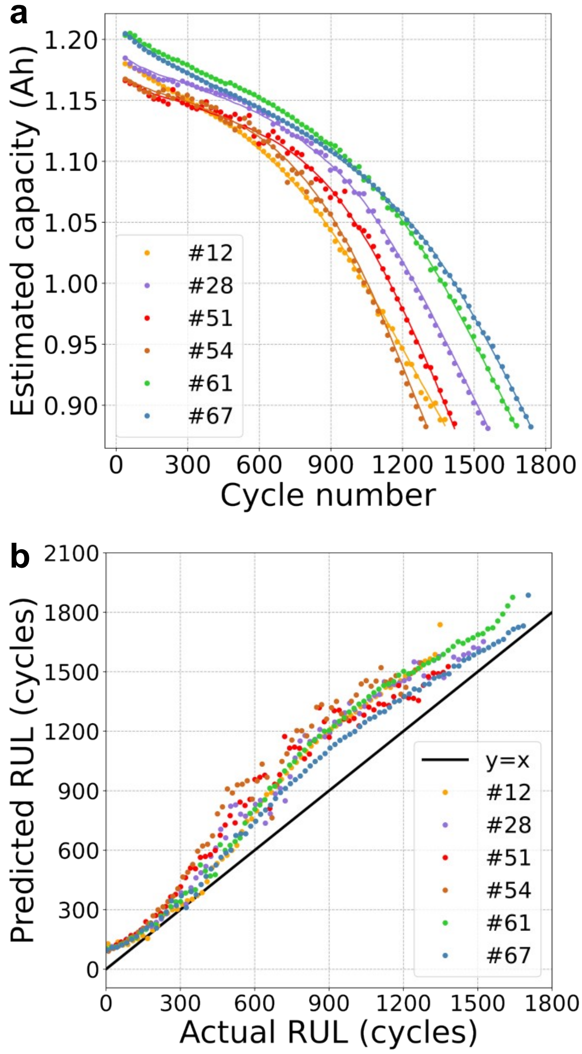

Fig. S13 Predicted health status versus actual health status for 6 testing protocols (protocols #12, #28, #51, #54, #61 and #67). The model is pre-trained by a public dataset with 22 NMC/graphite cells. (a) The predicted capacity versus actual capacity of the 6 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 6 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 6 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

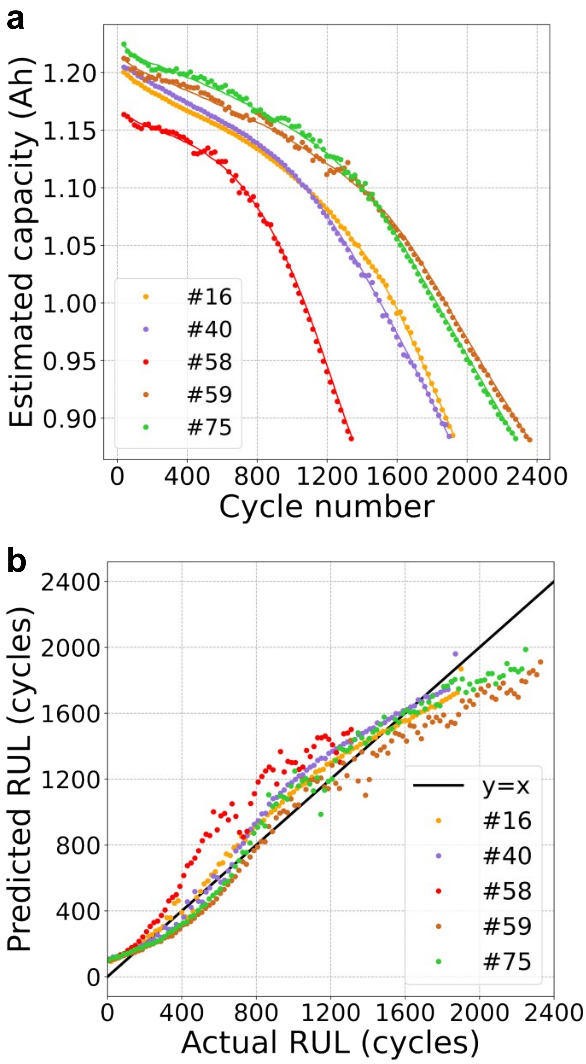

Fig. S14 Predicted health status versus actual health status for 5 testing protocols (protocols #16, #40, #58, #59 and #75). The model is pre-trained by a public dataset with 22 NMC/graphite cells. (a) The predicted capacity versus actual capacity of the 5 cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 5 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 5 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

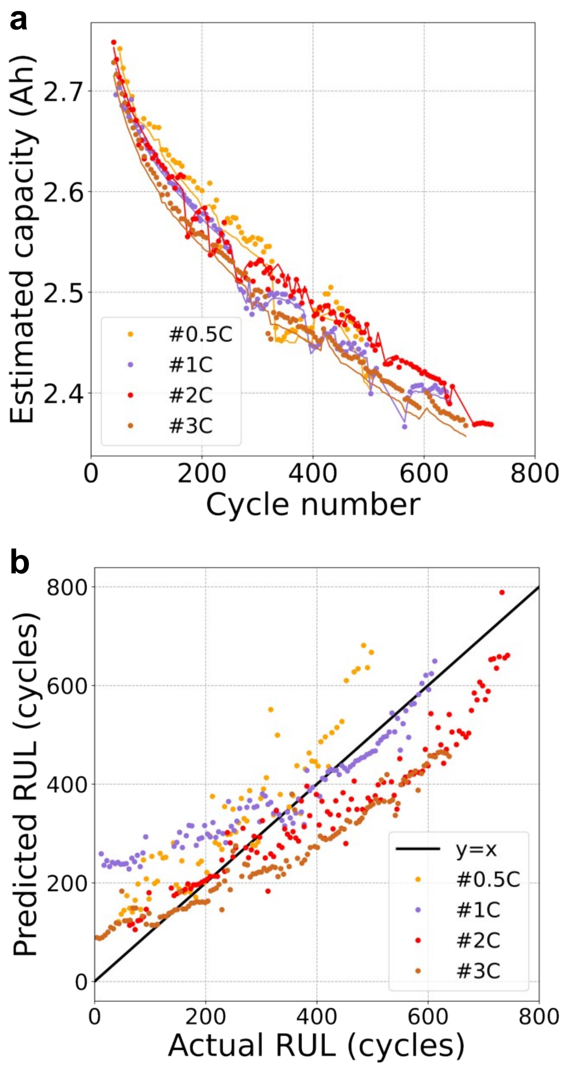

Fig. S15 Predicted health status versus actual health status for 4 testing protocols (CC discharge rates of 0.5C, 1C, 2C and 3C) in a public dataset with 22 NMC/graphite cells. Based on a cross-validation strategy, the model is pre-trained by the other discharge protocols of NMC/graphite cells when one protocol is used for testing. (a) The predicted capacity versus actual capacity of 4 testing cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 4 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 4 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

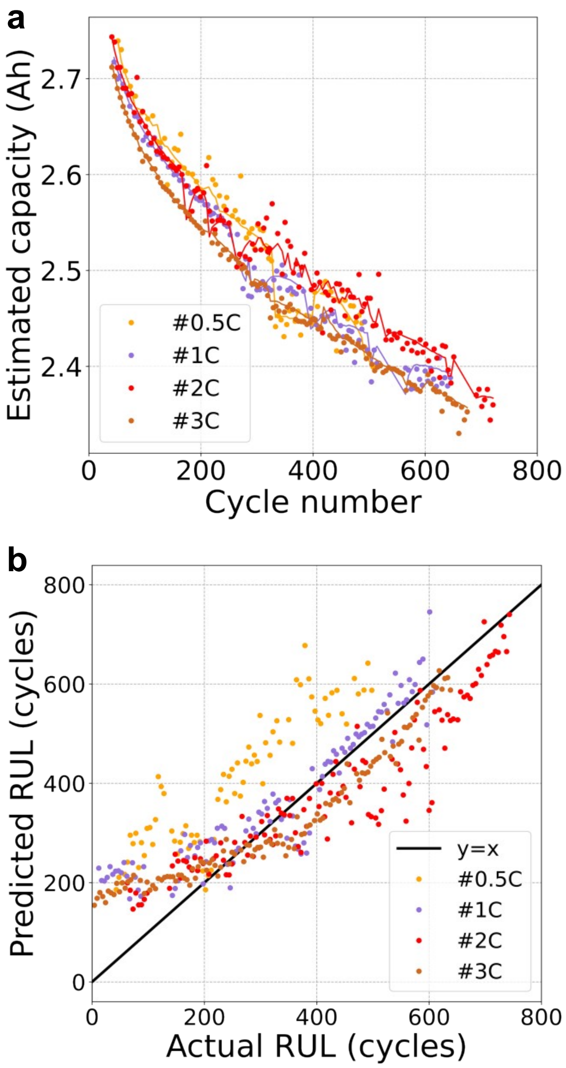

Fig. S16 Predicted health status versus actual health status for 4 testing protocols (CC discharge rates of 0.5C, 1C, 2C and 3C) in a public dataset with 22 NMC/graphite cells. The model is pre-trained by our dataset with 77 LFP/graphite cells. (a) The predicted capacity versus actual capacity of 4 testing cells, where the scatter points and solid lines represent the predicted capacities and the actual capacities, respectively. (b) The predicted RUL versus actual RUL of 4 testing cells, where the solid line represents the ideal result that the predicted RUL is always identical with the actual RUL and the scatter points with different colors represent the predicted RULs for 4 testing cells. For the convenience of showing the results, the predicted capacity and RUL are displayed every 20 cycles.

Supplemental Tables

Table S1 Description of 77 cells in our dataset. The cells were cycled with an identical charge protocol, but completely different discharge protocols (protocols #1-#77). Four-stage discharge from 100% SOC to a cut off voltage of 2 V. Diverse discharge protocols C1 (100% SOC-60% SOC)-C2 (60% SOC-40% SOC)-C3 (40% SOC-20% SOC)-C4 (20% SOC-2 V) are employed for different cells, where C1-C4 represent constant-current discharge rates at the four stages, respectively.

Protocol	**Channel**	**Data attribute**	**Cycle life (cycles)**	**C1 (100%-60%)**	**C2 (60%-40%)**	**C3 (40%-20%)**	**C4 (20%-2V)**
#1	1-1	Test	1,504	5C	1C	1C	1C
#2	1-2	Test	2,678	5C	1C	2C	1C
#3	1-3	Train	1,858	5C	1C	3C	1C
#4	1-4	Train	1,500	5C	1C	4C	1C
#5	1-5	Train	1,971	5C	1C	5C	1C
#6	1-6	Train	1,143	5C	2C	1C	1C
#7	1-7	Train	1,678	5C	2C	2C	1C
#8	1-8	Train	2,285	5C	2C	3C	1C
#9	2-2	Train	2,651	5C	2C	5C	1C
#10	2-3	Train	1,751	5C	3C	1C	1C
#11	2-4	Train	1,499	5C	3C	2C	1C
#12	2-5	Test	1,386	5C	3C	3C	1C
#13	2-6	Train	1,572	5C	3C	4C	1C
#14	2-7	Train	2,202	5C	3C	5C	1C
#15	2-8	Train	1,481	5C	4C	1C	1C
#16	3-1	Test	1,938	5C	4C	2C	1C
#17	3-2	Train	2,283	5C	4C	3C	1C
#18	3-3	Train	1,649	5C	4C	4C	1C
#19	3-4	Train	1,766	5C	4C	5C	1C
#20	3-5	Train	2,657	5C	5C	1C	1C
#21	3-6	Train	2,491	5C	5C	2C	1C
#22	3-7	Train	2,479	5C	5C	3C	1C
#23	3-8	Train	2,342	5C	5C	4C	1C
#24	4-1	Train	2,217	5C	5C	5C	1C
#25	4-2	Train	1,782	4C	1C	1C	1C
#26	4-3	Train	1,142	4C	1C	2C	1C
#27	4-4	Train	1,491	4C	1C	3C	1C
#28	4-5	Test	1,561	4C	1C	4C	1C
#29	4-6	Train	1,380	4C	1C	5C	1C
#30	4-7	Train	2,216	4C	2C	1C	1C
#31	4-8	Train	1,706	4C	2C	2C	1C
#32	5-1	Train	2,507	4C	2C	3C	1C
#33	5-2	Train	1,926	4C	2C	4C	1C

#34	5-3	Test	2,689	4C	2C	5C	1C
#35	5-4	Train	1,962	4C	3C	1C	1C
#36	5-5	Train	1,583	4C	3C	2C	1C
#37	5-6	Train	2,460	4C	3C	3C	1C
#38	5-7	Train	1,448	4C	3C	4C	1C
#39	6-1	Test	1,609	4C	4C	1C	1C
#40	6-2	Test	1,908	4C	4C	2C	1C
#41	6-3	Train	1,804	4C	4C	3C	1C
#42	6-4	Train	1,717	4C	4C	4C	1C
#43	6-5	Train	2,178	4C	4C	5C	1C
#44	6-6	Test	2,468	4C	5C	1C	1C
#45	6-8	Test	2,450	4C	5C	3C	1C
#46	7-1	Train	1,690	4C	5C	4C	1C
#47	7-2	Train	2,030	4C	5C	5C	1C
#48	7-3	Train	1,295	3C	1C	1C	1C
#49	7-4	Train	1,393	3C	1C	2C	1C
#50	7-5	Test	1,875	3C	1C	3C	1C
#51	7-6	Test	1,419	3C	1C	4C	1C
#52	7-7	Train	1,685	3C	1C	5C	1C
#53	7-8	Train	1,938	3C	2C	1C	1C
#54	8-1	Test	1,308	3C	2C	2C	1C
#55	8-2	Train	2,041	3C	2C	3C	1C
#56	8-3	Train	2,290	3C	2C	4C	1C
#57	8-4	Train	1,885	3C	2C	5C	1C
#58	8-5	Test	1,348	3C	3C	1C	1C
#59	8-6	Test	2,365	3C	3C	2C	1C
#60	8-7	Train	2,047	3C	3C	3C	1C
#61	8-8	Test	1,679	3C	3C	4C	1C
#62	9-1	Train	2,057	3C	3C	5C	1C
#63	9-2	Train	2,143	3C	4C	1C	1C
#64	9-3	Train	1,905	3C	4C	2C	1C
#65	9-4	Test	1,975	3C	4C	3C	1C
#66	9-5	Train	2,168	3C	4C	4C	1C
#67	9-6	Test	1,742	3C	4C	5C	1C
#68	9-7	Train	2,012	3C	5C	1C	1C
#69	9-8	Train	2,308	3C	5C	2C	1C
#70	10-1	Test	1,702	3C	5C	3C	1C
#71	10-2	Train	1,697	3C	5C	4C	1C
#72	10-3	Train	1,848	3C	5C	5C	1C
#73	10-4	Test	1,811	2C	4C	1C	1C
#74	10-5	Train	2,030	2C	5C	2C	1C

#75	10-6	Test	2,285	2C	3C	3C	1C
#76	10-7	Test	1,783	2C	2C	4C	1C
#77	10-8	Train	1,400	2C	1C	5C	1C

Table S2 Summary for the deep learning network configuration. 'k', 's' and 'p' in the CONV module represent kernel size, stride and padding size, respectively. 'in' and 'out' in the FC stand for input size and output size, respectively.

Module	Layer	Configurations	Shape
Input	---	---	(1, 4, 100, 10)
CONV	Convolutional	filters = 8, k = (3, 1), s = (2, 1), p = 0	(1, 8, 49, 10)
	Max-pooling	k = (2, 1), s = (2, 1), p = 0	(1, 8, 24, 10)
	Convolutional	filters = 16, k = (3, 1), s = (2, 1), p = 0	(1, 16, 11, 10)
	Max-pooling	k = (2, 1), s = (2, 1), p = 0	(1, 16, 5, 10)
	Convolutional	filters = 64, k = (3, 1), s = (2, 1), p = 0	(1, 64, 2, 10)
	Max-pooling	k = (2, 1), s = (2, 1), p = 0	(1, 64, 1, 10)
---	Reshape	---	(1, 10, 64)
RE	LSTM cell	hidden units = 64	(1, 10, 64)
	Linear	(in, out) = (64, 64)	(1, 10, 64)
	LSTM cell	hidden units = 64	(1, 10, 64)
	Linear	(in, out) = (64, 64)	(1, 10, 64)
FC_1	Linear	(in, out) = (64, 10)	(1, 10, 1)
FC_2	Linear	(in, out) = (10, 1)	(1, 1)

Table S3 Evaluation results of capacity estimation for all testing protocols. The estimation is performed among different discharge protocols in our dataset. ‘w/’ and ‘w/o’ represent the estimation results with model transfer and without model transfer, respectively.

Protocol	**Channel**	**RMSE (mAh)**		R^2		**MAPE (%)**	
		w/o	**w/**	**w/o**	**w/**	**w/o**	**w/**
#1	1-1	4.61	3.45	0.997	0.998	0.323	0.263
#2	1-2	3.70	3.47	0.998	0.998	0.241	0.209
#12	2-5	1.80	2.57	1.00	0.999	0.125	0.182
#16	3-1	3.93	1.45	0.998	1.00	0.282	0.102
#28	4-5	1.88	2.02	1.00	0.999	0.134	0.140
#34	5-3	1.18	2.92	1.00	0.999	0.0809	0.184
#39	6-1	3.11	3.74	0.999	0.998	0.224	0.268
#40	6-2	6.75	1.32	0.995	1.00	0.511	0.0892
#44	6-6	7.84	3.32	0.992	0.999	0.577	0.227
#45	6-8	2.49	2.21	0.999	0.999	0.170	0.139
#50	7-5	3.13	1.80	0.999	1.00	0.247	0.127
#51	7-6	2.37	2.64	0.999	0.999	0.172	0.182
#54	8-1	1.95	3.20	0.999	0.998	0.136	0.213
#58	8-5	1.52	4.49	1.00	0.997	0.102	0.328
#59	8-6	3.63	2.43	0.999	0.999	0.253	0.161
#61	8-8	2.76	2.18	0.999	0.999	0.175	0.133
#65	9-4	3.13	2.22	0.999	0.999	0.219	0.146
#67	9-6	2.61	1.37	0.999	1.00	0.190	0.0966
#70	10-1	1.49	1.91	1.00	0.999	0.107	0.139
#73	10-4	2.15	2.07	0.999	0.999	0.150	0.152
#75	10-6	6.75	3.12	0.995	0.999	0.484	0.202
#76	10-7	2.41	2.76	0.999	0.999	0.183	0.195
Mean	-	3.24	2.57	0.998	0.999	0.231	0.176

Table S4 Evaluation results of RUL prediction for all testing protocols. The prediction is performed among different discharge protocols in our dataset. 'Elastic net' represents RUL prediction using the elastic net[4] and manual features. 'w/' and 'w/o' represent the prediction results with model transfer and without model transfer, respectively.

Protocol	Channel	RMSE (cycles)			R^2			MAPE (%)		
		Elastic net	w/o	w/	Elastic net	w/o	w/	Elastic net	w/o	w/
#1	1-1	252	63.8	42.7	0.646	0.977	0.990	14.1	3.64	2.16
#2	1-2	722	262	272	0.102	0.882	0.873	23.2	8.64	8.52
#12	2-5	365	390	364	0.122	/	0.126	19.4	23.5	22.0
#16	3-1	422	104	133	0.407	0.964	0.941	17.0	4.43	6.00
#28	4-5	313	307	301	0.493	0.512	0.531	16.1	15.7	16.3
#34	5-3	757	279	301	0.0225	0.867	0.845	24.2	8.94	9.47
#39	6-1	310	147	120	0.532	0.896	0.930	14.3	7.69	6.01
#40	6-2	413	96.3	141	0.415	0.968	0.932	15.5	4.08	5.98
#44	6-6	672	400	349	0.0821	0.675	0.753	22.4	14.5	12.6
#45	6-8	609	334	359	0.236	0.769	0.734	20.3	12.6	13.6
#50	7-5	372	58.0	46.2	0.508	0.988	0.992	15.6	2.70	2.01
#51	7-6	372	295	263	0.129	0.453	0.566	21.5	15.7	15.6
#54	8-1	303	200	231	0.319	0.702	0.603	19.7	12.8	15.6
#58	8-5	281	45.3	57.1	0.449	0.986	0.977	17.7	2.64	3.30
#59	8-6	527	91.8	81.5	0.386	0.981	0.985	18.1	3.41	2.97
#61	8-8	412	363	382	0.245	0.411	0.349	18.2	18.2	19.3
#65	9-4	431	104	74.3	0.406	0.966	0.982	16.9	4.56	3.17
#67	9-6	403	297	292	0.328	0.634	0.649	18.4	15.4	15.1
#70	10-1	386	111	81.0	0.355	0.947	0.972	17.6	5.03	4.24
#73	10-4	331	67.5	60.9	0.582	0.983	0.986	14.3	3.15	2.95
#75	10-6	485	112	84.0	0.442	0.970	0.983	16.0	3.93	3.12
#76	10-7	405	86.7	52.8	0.353	0.970	0.989	16.7	3.25	1.97
Mean	-	434	192	186	0.344	0.795	0.804	18.1	8.84	8.72

Table S5 Evaluation results of capacity estimation for all testing protocols. The model is pre-trained by LFP/graphite cells in a public dataset and is transferred to our testing cells for capacity estimation. 'w/' and 'w/o' represent the estimation results with model transfer and without model transfer, respectively.

Protocol	Channel	RMSE (mAh)		R^2		MAPE (%)	
		w/o	w/	w/o	w/	w/o	w/
#1	1-1	96.7	5.48	/	0.995	6.97	0.401
#2	1-2	112	4.58	/	0.997	7.97	0.335
#12	2-5	114	4.07	/	0.998	9.47	0.300
#16	3-1	54.3	4.33	0.588	0.997	3.82	0.294
#28	4-5	77.1	5.45	0.152	0.996	5.78	0.390
#34	5-3	109	4.72	/	0.997	7.78	0.340
#39	6-1	107	5.08	/	0.996	7.76	0.375
#40	6-2	109	4.16	/	0.998	8.80	0.288
#44	6-6	134	4.24	/	0.998	10.9	0.301
#45	6-8	77.0	3.97	0.250	0.998	5.40	0.281
#50	7-5	64.0	4.98	0.396	0.996	4.77	0.357
#51	7-6	70.6	4.75	0.164	0.996	5.37	0.313
#54	8-1	86.1	6.25	/	0.994	6.53	0.410
#58	8-5	89.6	6.48	/	0.993	6.82	0.444
#59	8-6	110	4.98	/	0.997	8.27	0.362
#61	8-8	66.5	4.21	0.414	0.998	4.87	0.291
#65	9-4	55.8	4.46	0.577	0.997	4.08	0.312
#67	9-6	48.3	3.88	0.670	0.998	3.23	0.259
#70	10-1	55.2	2.64	0.496	0.999	4.25	0.193
#73	10-4	66.6	5.27	0.369	0.996	4.92	0.378
#75	10-6	100	4.55	/	0.998	7.60	0.321
#76	10-7	74.0	3.91	0.250	0.998	5.69	0.270
Mean	-	85.3	4.65	0.197	0.997	6.41	0.328

Table S6 Evaluation results of RUL prediction for all testing protocols. The model is pre-trained by LFP/graphite cells in a public dataset and is transferred to our testing cells for capacity estimation. 'Elastic net' represents RUL prediction using the elastic net[4] and manual features. 'w/' and 'w/o' represent the prediction results with model transfer and without model transfer, respectively.

Protocol	Channel	RMSE (cycles)			R^2			MAPE (%)		
		Elastic net	w/o	w/	Elastic net	w/o	w/	Elastic net	w/o	w/
#1	1-1	451	503	110	/	/	0.925	27.7	29.6	5.60
#2	1-2	944	730	663	/	/	0.201	28.7	24.3	21.7
#12	2-5	479	533	85.6	/	/	0.946	31.0	34.9	5.16
#16	3-1	599	841	232	/	/	0.808	26.7	35.6	10.4
#28	4-5	497	760	76.9	/	/	0.966	28.5	36.8	3.72
#34	5-3	951	716	683	/	/	0.159	28.8	23.8	22.0
#39	6-1	496	642	181	/	/	0.825	27.5	35.2	9.73
#40	6-2	610	596	185	/	/	0.873	28.3	28.7	8.42
#44	6-6	840	1,133	523	/	/	0.411	27.9	41.0	18.8
#45	6-8	848	701	595	/	/	0.225	29.1	25.9	21.6
#50	7-5	576	802	187	/	/	0.865	26.5	33.1	8.31
#51	7-6	439	732	61.5	/	/	0.974	26.5	37.2	3.63
#54	8-1	383	644	73.9	/	/	0.954	25.9	33.1	4.64
#58	8-5	360	532	52.1	/	/	0.979	24.8	26.8	3.32
#59	8-6	817	615	400	/	/	0.624	29.2	22.9	15.0
#61	8-8	521	1,131	75.8	/	/	0.972	27.2	55.6	3.58
#65	9-4	614	866	205	/	/	0.855	26.5	34.9	9.07
#67	9-6	521	1,151	143	/	/	0.908	24.9	54.1	7.19
#70	10-1	539	687	176	/	/	0.853	27.9	31.4	8.88
#73	10-4	521	795	152	/	/	0.905	24.9	34.9	6.64
#75	10-6	781	767	285	/	/	0.795	30.1	29.5	11.4
#76	10-7	559	846	141	/	/	0.915	27.9	38.3	6.80
Mean	-	607	760	240	/	/	0.770	27.6	34.0	9.80

Table S7 Evaluation results of capacity estimation for all testing protocols. The model is pre-trained by a public dataset with 22 NMC/graphite cells and is transferred to our testing cells for capacity estimation. 'w/' and 'w/o' represent the estimation results with model transfer and without model transfer, respectively.

Protocol	Channel	RMSE (mAh)		R^2		MAPE (%)	
		w/o	w/	w/o	w/	w/o	w/
#1	1-1	80.2	3.30	/	0.998	5.45	0.223
#2	1-2	67.0	3.20	0.437	0.999	4.27	0.206
#12	2-5	120	3.45	/	0.998	8.00	0.216
#16	3-1	94.6	2.18	/	0.999	6.03	0.135
#28	4-5	82.2	3.58	/	0.998	5.70	0.239
#34	5-3	75.1	2.97	0.229	0.999	4.31	0.183
#39	6-1	75.0	3.46	0.161	0.998	5.20	0.222
#40	6-2	108	2.51	/	0.999	6.78	0.146
#44	6-6	115	3.36	/	0.999	7.44	0.184
#45	6-8	87.5	3.08	/	0.999	5.88	0.187
#50	7-5	85.0	2.51	/	0.999	5.57	0.141
#51	7-6	97.7	3.94	/	0.997	6.91	0.267
#54	8-1	96.3	4.27	/	0.997	6.97	0.274
#58	8-5	97.3	4.13	/	0.997	7.12	0.273
#59	8-6	88.7	3.83	0.110	0.998	5.80	0.232
#61	8-8	106	1.91	/	1.00	6.43	0.116
#65	9-4	99.6	2.44	/	0.999	6.08	0.141
#67	9-6	91.7	2.06	/	0.999	5.85	0.131
#70	10-1	95.6	3.34	/	0.998	6.74	0.225
#73	10-4	85.8	2.46	/	0.999	5.51	0.136
#75	10-6	81.7	3.97	0.308	0.998	5.51	0.247
#76	10-7	103	1.91	/	1.00	6.36	0.118
Mean	-	92.4	3.08	0.249	0.999	6.09	0.193

Table S8 Evaluation results of RUL prediction for all testing protocols. The model is pre-trained by a public dataset with 22 NMC/graphite cells and is transferred to our testing cells for capacity estimation. 'Elastic net' represents RUL prediction using the elastic net[4] and manual features. 'w/' and 'w/o' represent the prediction results with model transfer and without model transfer, respectively.

Protocol	Channel	RMSE (cycles)			R^2			MAPE (%)		
		Elastic net	w/o	w/	Elastic net	w/o	w/	Elastic net	w/o	w/
#1	1-1	533	894	164	/	/	0.850	29.1	48.9	9.16
#2	1-2	1,200	1,110	470	/	/	0.620	37.1	32.0	14.9
#12	2-5	474	931	221	/	/	0.677	28.2	58.8	13.7
#16	3-1	779	449	84.8	/	/	0.976	32.3	18.6	3.79
#28	4-5	562	939	227	/	/	0.732	29.1	49.9	12.8
#34	5-3	1,210	1,190	492	/	/	0.587	37.2	34.6	15.3
#39	6-1	595	798	87.3	/	/	0.963	30.5	40.5	4.05
#40	6-2	758	778	104	/	/	0.963	32.5	30.7	4.30
#44	6-6	1,080	844	370	/	/	0.722	36.2	28.0	12.9
#45	6-8	1,070	1,080	384	/	/	0.696	35.8	34.7	13.7
#50	7-5	738	621	124	/	/	0.945	31.7	26.7	5.56
#51	7-6	486	1,080	246	/	/	0.620	27.7	62.8	15.4
#54	8-1	426	1,030	308	/	/	0.293	26.6	63.2	21.2
#58	8-5	448	981	265	/	/	0.511	27.3	58.9	17.4
#59	8-6	1,020	941	197	/	/	0.914	35.4	33.0	6.03
#61	8-8	629	496	220	/	/	0.784	30.2	22.2	11.7
#65	9-4	796	627	106	/	/	0.964	32.5	23.6	4.60
#67	9-6	666	529	145	/	/	0.913	30.9	25.8	7.53
#70	10-1	643	1,030	111	/	/	0.947	30.4	51.1	5.55
#73	10-4	703	543	158	/	/	0.905	31.3	24.0	7.37
#75	10-6	972	948	131	/	/	0.959	34.9	34.5	4.37
#76	10-7	688	476	139	/	/	0.924	31.2	20.6	6.64
Mean	-	749	832	216	/	/	0.794	31.7	37.4	9.90

Table S9 The evaluation metrics for battery health status prediction by fine-tuning the convolutional module at the model transfer stage.

Capacity	**RMSE (mAh)**	R^2	**MAPE (%)**
Task A	2.67	0.999	0.180
Task B	8.36	0.990	0.563
Task C	10.90	0.984	0.741
RUL	**RMSE (cycles)**	R^2	**MAPE (%)**
Task A	210	0.781	9.66
Task B	344	0.565	13.50
Task C	246	0.777	9.66

Supplemental Notes

Note S1 Online fine-tuning stage of deep transfer learning model

To predict the health status at any cycle of each discharge protocol, a fine-tuning stage is used to online fine-tune the pre-trained deep learning model. We use input feature curves and observed discharge capacities in recent 30 charge-discharge cycles (actually, 10 charge-discharge cycles after a resample) to fine-tune the recurrent module and predict the health status. In real applications, only the discharge capacities at the first 9 charge-discharge cycles can be obtained by Ampere hour counting, but the capacity at the current charge-discharge cycle and RULs of all 10 charge-discharge cycles are unknown. Hence, we employ the input feature curves of all 10 charge-discharge cycles and observed discharge capacities at the first 9 charge-discharge cycles to fine-tune the model weights. In detail, random d of 9 observed discharge capacities are used to fine-tune the weights in the recurrent module and discharge capacities of the other 9-d charge-discharge cycles are used to validate the performance. For example, α = [1, 1, 0, 0, 0, 0, 0, 0, 0] means the discharge capacities at the first 2 charge-discharge cycles are used to fine-tune the model weights and discharge capacities at the last 7 charge-discharge cycles are used to validate the model performance. Here, an early-stopping technique is applied to guarantee the generalization of the fine-tuned model. Then, the discharge capacities at the current charge-discharge cycle and the previous 9 charge-discharge cycles can be estimated and updated using the fine-tuned model. Based on the estimated capacities, the RUL at the current charge-discharge cycle can be online predicted using the FC_2 module.

Through the above online fine-tuning process, the health status at any charge-discharge cycle of any discharge protocol could be predicted using cycling data in recent 30 charge-discharge cycles, which satisfies the real-time and personalized requirements.

Note S2 The benchmark models

Two benchmark models are used for performance comparison: a feature-based machine learning method and an individual deep learning method without a model transfer stage.

A feature-based machine learning approach is used as the first benchmark model, in which the formulae for the 6 selected features[3] based on the charging voltage-vs.-capacity curves are:

- Mean of square = $\log\left|\frac{1}{p-1}\sum_{i=1}^{p}(\Delta V(Q))^2\right|$
- Minimum = $\log\left|min(\Delta V(Q))\right|$
- Maximum = $\log\left|max(\Delta V(Q))\right|$
- Mean = $\log\left|\bar{\Delta V}(Q)\right|$
- Skewness = $\log\left|\frac{\frac{1}{p}\sum_{i=1}^{p}(\Delta V(Q)-\bar{\Delta V}(Q))^3}{\left(\sqrt{\frac{1}{p}\sum_{i=1}^{p}(\Delta V(Q)-\bar{\Delta V}(Q))^2}\right)^3}\right|$
- Kurtosis = $\log\left|\frac{\frac{1}{p}\sum_{i=1}^{p}(\Delta V(Q)-\bar{\Delta V}(Q))^4}{\left(\frac{1}{p}\sum_{i=1}^{p}(\Delta V(Q)-\bar{\Delta V}(Q))^2\right)^2}\right|$

where log indicates log base *e*. These features are functions of

$\Delta V(Q) = V_i(Q) - V_j(Q), \Delta V(Q) \in R^p$

$$\bar{\Delta V}(Q) = \frac{1}{p}\sum_{i=1}^{p}\Delta V(Q)$$

where the subscripts i and j indicate the cycle number and the range of Q is from 0 Ah to the min{ Q_i^{end}, Q_j^{end} }, where Q_i^{end} and Q_j^{end} are the capacities at the end of charge process for cycles i and j, respectively. In this work, the i^{th} and j^{th} cycles are the 30^{th} cycle (i.e., the current cycle) and the 2^{nd} cycle in a total of 30 cycles of each input sample, respectively. A popular regularization technique called the elastic net is used as our benchmark model with the manually selected features to predict the RULs at all charge-discharge cycles of testing cells.

In addition, an individual deep learning method without a model transfer stage is also employed as a benchmark model for performance comparison. We pre-train the model using 55 training cells and directly predict the capacities and RULs of the 22 testing cells without the fine-tuning stage.

Supplemental References

1 P. M. Attia, A. Grover, N. Jin, K. A. Severson, T. M. Markov, Y. H. Liao, M. H. Chen, B. Cheong, N. Perkins, Z. Yang, P. K. Herring, M. Aykol, S. J. Harris, R. D. Braatz, S. Ermon and W. C. Chueh, *Nature*, 2020, **578**, 397–402.

2 Y. Preger, H. M. Barkholtz, A. Fresquez, D. L. Campbell, B. W. Juba, J. Romàn-Kustas, S. R. Ferreira and B. Chalamala, *J. Electrochem. Soc.*, 2020, **167**, 120532.

3 B. Jiang, W. E. Gent, F. Mohr, S. Das, M. D. Berliner, M. Forsuelo, H. Zhao, P. M. Attia, A. Grover, P. K. Herring, M. Z. Bazant, S. J. Harris, S. Ermon, W. C. Chueh and R. D. Braatz, *Joule*, 2021, **5**, 3187–3203.

4 H. Zou and T. Hastie, *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 2005, **67**, 301–320.