Supplementary information

A data-driven XRD analysis protocol for phase identification and phase-fraction prediction of multiphase inorganic compounds

Jin-Woong Lee,[a,†] Woon Bae Park,[b,†] Minseuk Kim,[a] Satendra Pal Singh,[a] Myoungho Pyo,[b,*] and Kee-Sun Sohn[a,*]

[a]Nanotechnology & Advanced Materials Engineering, Sejong University, 209 Neungdong-ro, Gwangjin-gu, Seoul, 143-747, South Korea

[b]Department of Printed Electronics, Sunchon National University, 291-19 Jungang-ro, Sunchon, Chonnam, 540-742, South Korea

[†]These authors contributed equally

Corresponding authors: kssohn@sejong.ac.kr, mho@sunchon.ac.kr

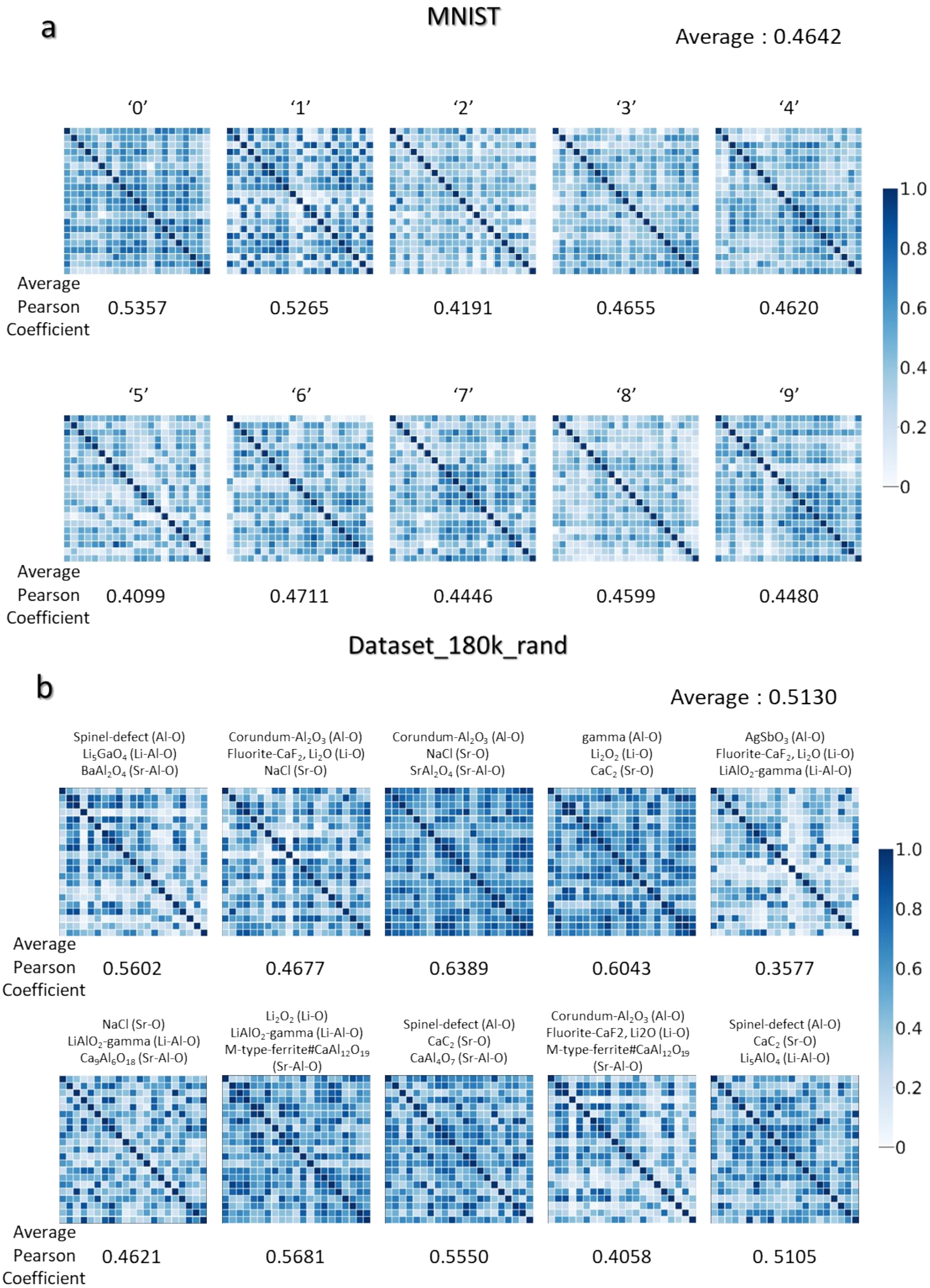
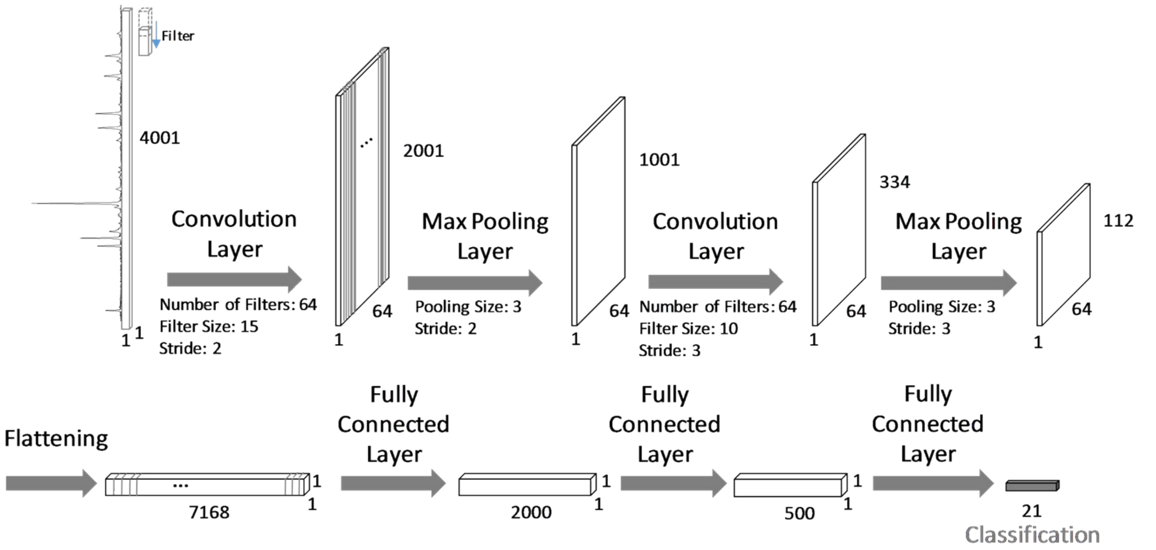
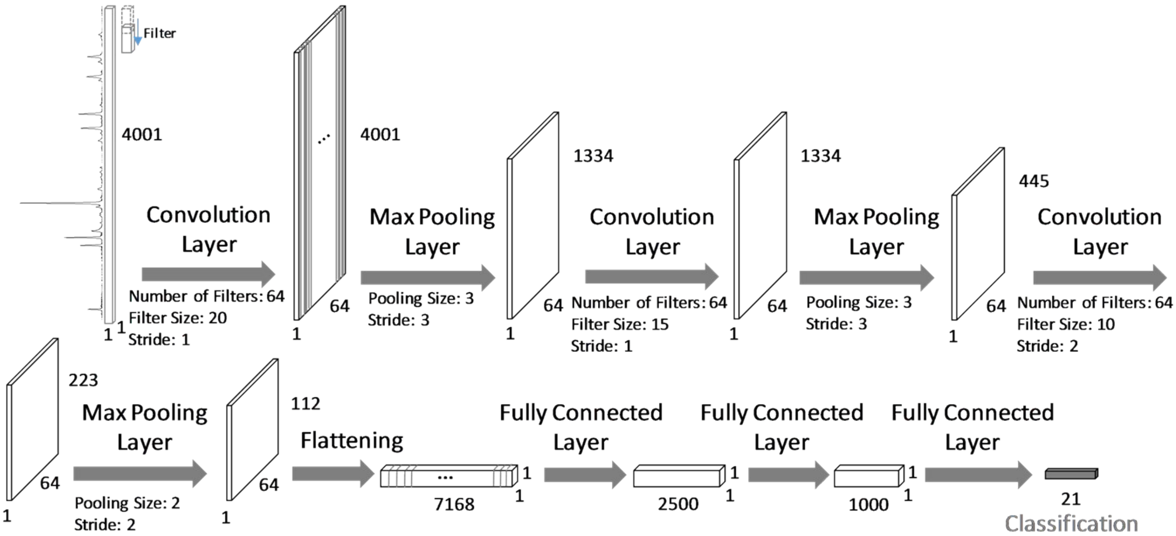


Figure S1 The similarity issue should be quantized by a reliable correlation metric such as the Pearson correlation coefficient. The correlation among the samples in a class (so-called intra-class correlation) was quantized using the Pearson correlation matrix. **a** Pearson correlation coefficient matrixes for 10 randomly chosen classes (in fact, including the classes that all the real-world data belong to) were obtained from the D2 dataset. A randomly chosen 66 samples from each of the 10 classes were used for constituting the following Pearson correlation coefficient matrixes. **b** The well-known MNIST datasets were also analyzed in terms of the correlation for comparison to the D2 dataset. Randomly chosen 21 samples from each of the classes (10 digits) were used to constitute the following Pearson correlation coefficient matrixes. **c** Pearson correlation coefficient matrixes obtained from the Dataset_180k_rand in our previous report[2].

CNN_2 Architecture



CNN_3 Architecture



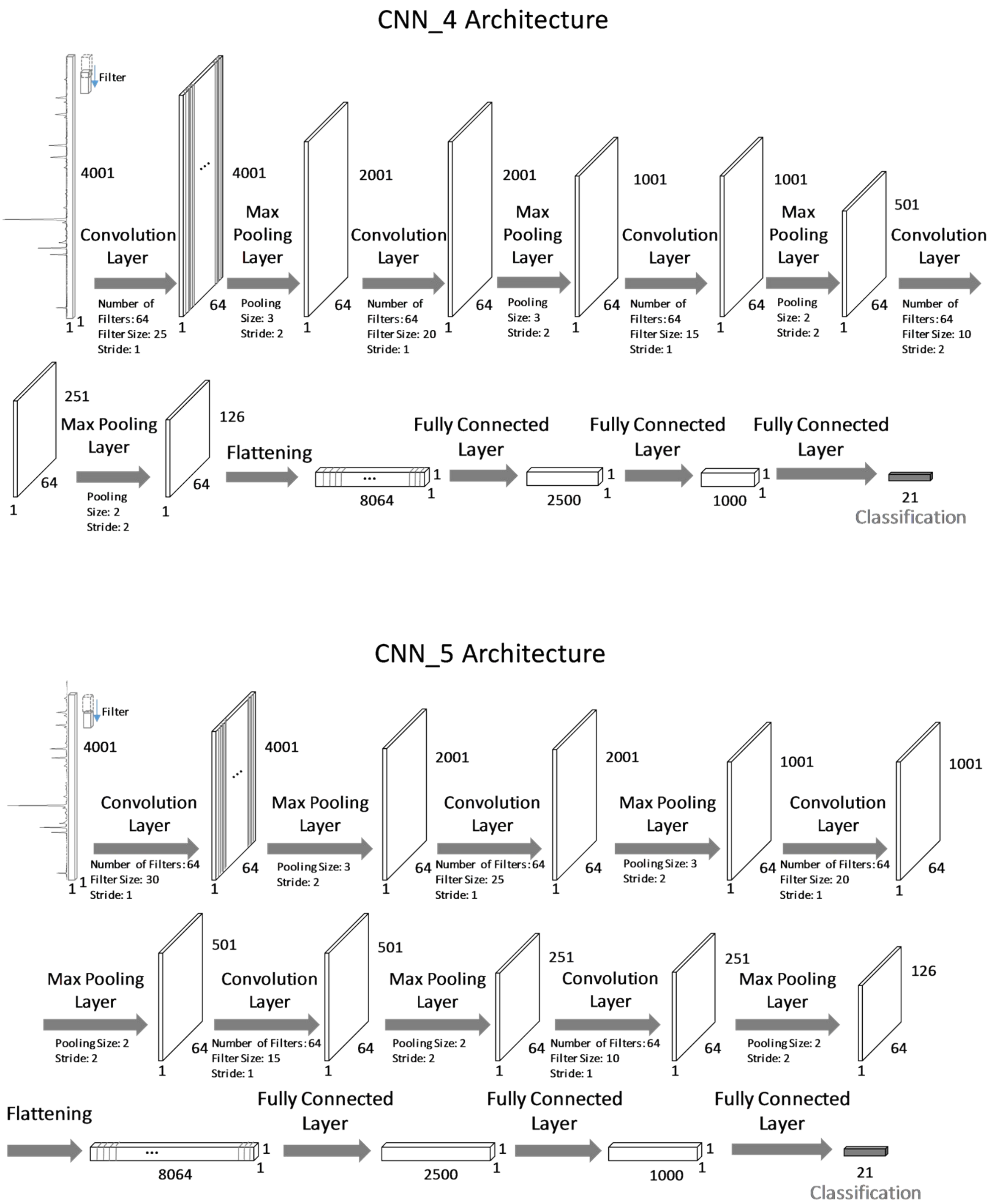


Figure S2 CNN_2 ~ CNN_5 architectures consisting of two ~ five convolution/max pooling layers, respectively. Three ensuing fully connected common layers followed. The number of filters, the kernel size, the pooling size, and the stride are also given. The padding type was the "SAME" for all CNNs. The drop-out was adopted only for the fully connected layers.

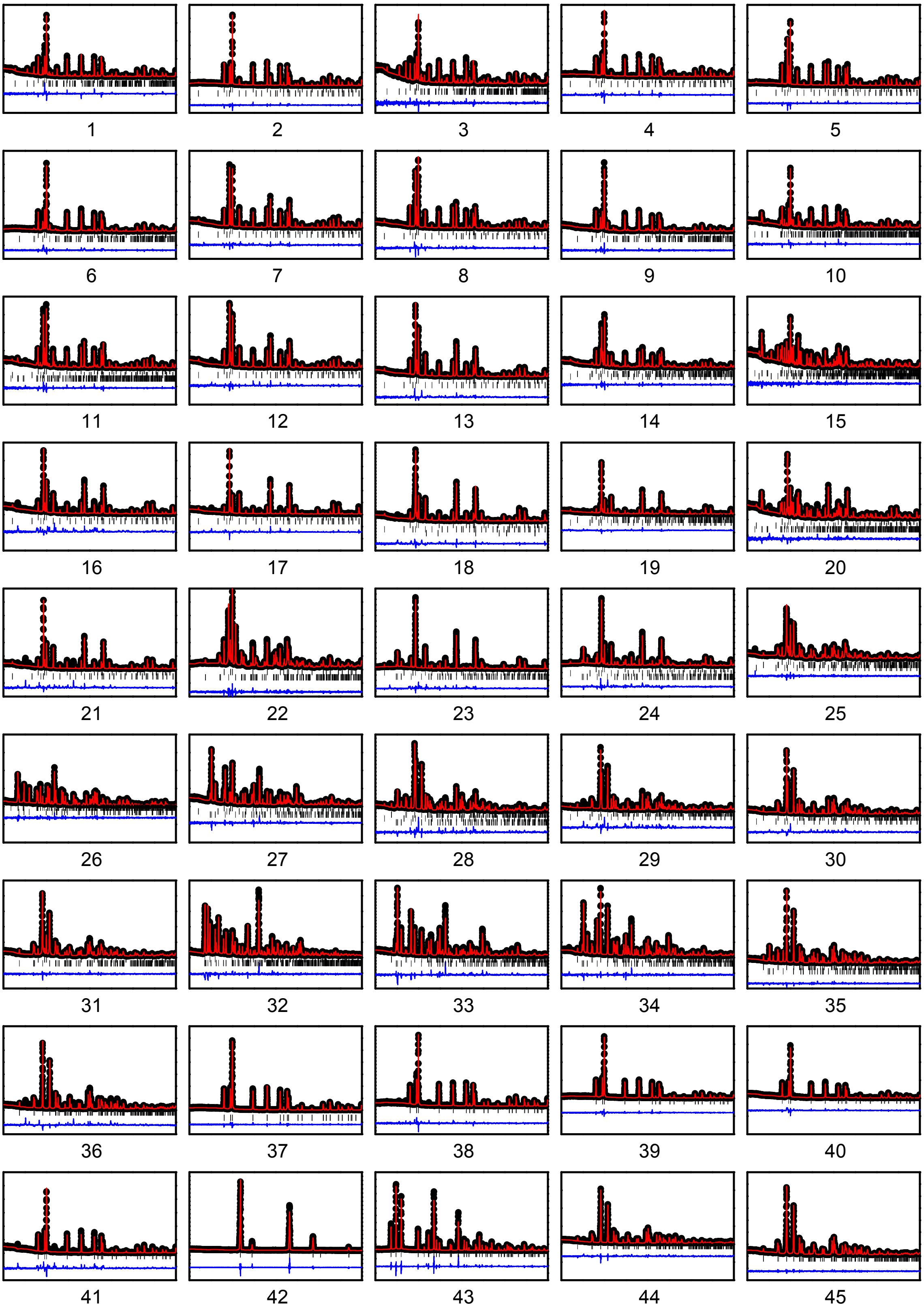


Figure S3 The Rietveld refinement results for 45 real-world XRD patterns.

The list of 218 constituent compounds (21 structure groups)

21 Groups	ICSD_ Num	Compound	Space Group	Lattice Parameter a	b	c	α	β	γ
Fluorite-CaF2, Li_2O (Li-O) Srtuc._Num_0	22402	Li_2O	Fm-3m (225)	4.61	4.61	4.61	90	90	90
	54368	Li_2O	Fm-3m (225)	4.610(5)	4.610(5)	4.610(5)	90	90	90
	57411	Li_2O	Fm-3m (225)	4.623	4.623	4.623	90	90	90
	60431	Li_2O	Fm-3m (225)	4.628	4.628	4.628	90	90	90
	173180	Li_2O	Fm-3m (225)	4.614(1)	4.614(1)	4.614(1)	90	90	90
	173193	Li_2O	Fm-3m (225)	4.6124(1)	4.6124(1)	4.6124(1)	90	90	90
	173206	Li_2O	Fm-3m (225)	4.6128(4)	4.6128(4)	4.6128(4)	90	90	90
	182024	Li_2O	Fm-3m (225)	4.728(5)	4.728(5)	4.728(5)	90	90	90
	182025	Li_2O	Fm-3m (225)	4.764(5)	4.764(5)	4.764(5)	90	90	90
	182026	Li_2O	Fm-3m (225)	4.782(5)	4.782(5)	4.782(5)	90	90	90
	182027	Li_2O	Fm-3m (225)	4.807(5)	4.807(5)	4.807(5)	90	90	90
	182028	Li_2O	Fm-3m (225)	4.837(5)	4.837(5)	4.837(5)	90	90	90
	257372	Li_2O	Fm-3m (225)	4.61549(5)	4.61549(5)	4.61549(5)	90	90	90
	642216	Li_2O	Fm-3m (225)	4.693(5)	4.693(5)	4.693(5)	90	90	90
LiO_2 (Li-O) Srtuc._Num_1	25530	Li_2O_2	$P6_3/mmc$ (194)	3.142	3.142	7.65	90	90	120
No structure type (Li-O) Srtuc._Num_2	24143	Li_2O_2	P-6 (174)	6.305	6.305	7.71	90	90	120
No structure type (Li-O) Srtuc._Num_3	108886	Li_2O	R-3mH (166)	3.624	3.624	7.97	90	90	120
La_2O_3 Srtuc._Num_4	24693	La_2O_3	P-3m1 (164)	3.94	3.94	6.13	90	90	120
	28555	La_2O_3	$P6_3/mmc$ (194)	3.9373	3.9373	6.1299	90	90	120
	56771	La_2O_3	P-3m1 (164)	3.934(1)	3.934	6.136(2)	90	90	120
	100204	La_2O_3	P-3m1 (164)	3.9381(3)	3.9381(3)	6.1361(6)	90	90	120
	100205	La_2O_3	P-3m1 (164)	4.039(2)	4.039(2)	6.403(3)	90	90	120
	100208	La_2O_3	$P6_3/mmc$ (194)	4.057(2)	4.057(2)	6.430(3)	90	90	120
	100209	La_2O_3	$P6_3/mmc$ (194)	4.057(2)	4.057(2)	6.430(3)	90	90	120
	100210	La_2O_3	$P6_3/mmc$ (194)	4.057(2)	4.057(2)	6.430(3)	90	90	120
	100215	La_2O_3	$P6_3/mmc$ (194)	4.057(2)	4.057(2)	6.430(3)	90	90	120
	192270	La_2O_3	P-3m1 (164)	3.9137(7)	3.9137(7)	6.0984(0)	90	90	120
	257585	La_2O_3	P-3m1 (164)	3.968(1)	3.968(1)	6.215(1)	90	90	120
	641599	La_2O_3	P-3m1 (164)	3.936	3.936	6.128	90	90	120
	641603	La_2O_3	P-3m1 (164)	3.937	3.937	6.13	90	90	120
La_2O_3(cI5) Srtuc._Num_5	44692	La_2O_3	Im-3m (229)	4.51(1)	4.51	4.51	90	90	90
Zirconia-ZrO_2(HT) Srtuc._Num_6	9993	ZrO_2	$P4_2/nmc$ S (137)	3.64	3.64	5.27	90	90	90
	23928	ZrO_2	$P4_2/nmc$ S (137)	3.64	3.64	5.27	90	90	90
	51051	ZrO_2	$P4_2/nmc$ Z (137)	3.64	3.64	5.27	90	90	90
	66781	ZrO_2	$P4_2/nmc$ Z (137)	3.5960(1)	3.5960(1)	5.1841(2)	90	90	90
	66782	ZrO_2	$P4_2/nmc$ Z (137)	3.5958(1)	3.5958(1)	5.1844(2)	90	90	90
	66783	ZrO_2	$P4_2/nmc$ Z (137)	3.5957(8)	3.5957(8)	5.1844(2)	90	90	90
	66784	ZrO_2	$P4_2/nmc$ Z (137)	3.5957(1)	3.5957(1)	5.1845(2)	90	90	90
	66785	ZrO_2	$P4_2/nmc$ Z	3.5961(1)	3.5961(1)	5.1843(2)	90	90	90

			(137)						
	66786	ZrO_2	$P4_2/nmc$ Z (137)	3.5958(1)	3.5958(1)	5.1849(2)	90	90	90
	66787	ZrO_2	$P4_2/nmc$ Z (137)	3.5957(2)	3.5957(2)	5.1850(3)	90	90	90
	66788	ZrO_2	$P4_2/nmc$ Z (137)	3.5936(4)	3.5936(4)	5.1814(6)	90	90	90
	66789	ZrO_2	$P4_2/nmc$ Z (137)	3.5916(8)	3.5916(8)	5.179(2)	90	90	90
	68589	ZrO_2	$P4_2/nmc$ S (137)	3.5961(2)	3.5961(2)	5.1770(4)	90	90	90
	68781	ZrO_2	$P4_2/nmc$ Z (137)	3.5925(3)	3.5925(3)	5.1837(5)	90	90	90
	70014	$ZrO_{1.99}$	$P4_2/nmc$ Z (137)	3.6067(4)	3.6067(4)	5.1290(7)	90	90	90
	70015	$ZrO_{1.88}$	$P4_2/nmc$ Z (137)	3.6067(4)	3.6067(4)	5.1520(8)	90	90	90
	72699	$Zr_{0.94}O_2$	$P4_2/nmc$ Z (137)	3.6231(11)	3.6231(11)	5.2412(28)	90	90	90
	72700	$Zr_{0.94}O_2$	$P4_2/nmc$ Z (137)	3.62802(6)	3.62802(6)	5.2493(14)	90	90	90
	72701	$Zr_{0.94}O_2$	$P4_2/nmc$ Z (137)	3.63187(12)	3.63187(12)	5.25943(26)	90	90	90
	72702	$Zr_{0.946}O_2$	$P4_2/nmc$ Z (137)	3.63981(11)	3.63981(11)	5.26166(15)	90	90	90
	72703	$Zr_{0.942}O_2$	$P4_2/nmc$ Z (137)	3.63273(7)	3.63273(7)	5.26189(12)	90	90	90
	72704	$Zr_{0.952}O_2$	$P4_2/nmc$ Z (137)	3.63700(9)	3.63700(9)	5.26957(17)	90	90	90
	72705	$Zr_{0.958}O_2$	$P4_2/nmc$ Z (137)	3.63981(10)	3.63981(10)	5.27454(18)	90	90	90
	72706	$Zr_{0.95}O_2$	$P4_2/nmc$ Z (137)	3.64370(10)	3.64370(10)	5.28172(19)	90	90	90
	72707	$Zr_{0.952}O_2$	$P4_2/nmc$ Z (137)	3.64548(11)	3.64548(11)	5.28497(20)	90	90	90
	72708	$Zr_{0.952}O_2$	$P4_2/nmc$ Z (137)	3.65762(13)	3.65762(13)	5.30475(24)	90	90	90
	72949	$ZrO_{1.95}$	$P4_2/nmc$ Z (137)	3.6067(4)	3.6067(4)	5.1758(8)	90	90	90
	72950	$ZrO_{1.99}$	$P4_2/nmc$ Z (137)	3.6136(3)	3.6136(3)	5.1909(5)	90	90	90
	72951	$ZrO_{1.96}$	$P4_2/nmc$ Z (137)	3.6227(3)	3.6227(3)	5.2056(6)	90	90	90
	72952	$ZrO_{1.97}$	$P4_2/nmc$ Z (137)	3.6358(4)	3.6358(4)	5.2257(7)	90	90	90
	72953	$ZrO_{1.98}$	$P4_2/nmc$ Z (137)	3.6411(4)	3.6411(4)	5.2340(7)	90	90	90
	72954	$ZrO_{1.96}$	$P4_2/nmc$ Z (137)	3.6499(4)	3.6499(4)	5.2483(8)	90	90	90
	75063	$(La_{0.1} Zr_{0.9}) O_{1.95}$	$P4_2/nmc$ Z (137)	3.6447(5)	3.6447(5)	5.2034(10)	90	90	90
	81036	$(Zr_{0.9} La_{0.1}) O_{1.95}$	$P4_2/nmc$ Z (137)	3.6377(7)	3.6377(7)	5.2059(26)	90	90	90
	85322	ZrO_2	$P4_2/nmc$ Z (137)	3.5984(5)	3.5984(5)	5.152(1)	90	90	90
	88022	ZrO_2	$P4_2/nmc$ Z (137)	3.6292(3)	3.6292(3)	5.1973(9)	90	90	90
	92090	ZrO_2	$P4_2/nmc$ Z	3.5948(2)	3.5948(2)	5.1824(7)	90	90	90

			(137)						
	92091	ZrO_2	$P4_2/nmc$ Z (137)	3.565(2)	3.565(2)	5.037(13)	90	90	90
	92092	ZrO_2	$P4_2/nmc$ Z (137)	3.512(3)	3.512(3)	4.988(8)	90	90	90
	92093	ZrO_2	$P4_2/nmc$ Z (137)	3.495(1)	3.495(1)	4.952(3)	90	90	90
	93028	ZrO_2	$P4_2/nmc$ S (137)	3.5815(12)	3.5815(12)	5.1685(40)	90	90	90
	93029	ZrO_2	$P4_2/nmc$ S (137)	3.5770(12)	3.5770(12)	5.1589(37)	90	90	90
	93030	ZrO_2	$P4_2/nmc$ S (137)	3.5781(3)	3.5781(3)	5.1623(8)	90	90	90
	93031	ZrO_2	$P4_2/nmc$ S (137)	3.5742(3)	3.5742(3)	5.1540(8)	90	90	90
	93032	ZrO_2	$P4_2/nmc$ S (137)	3.5794(2)	3.5794(2)	5.1647(7)	90	90	90
	93033	ZrO_2	$P4_2/nmc$ S (137)	3.5747(2)	3.5747(2)	5.1558(7)	90	90	90
	93123	ZrO_2	$P4_2/nmc$ S (137)	3.612(1)	3.612(1)	5.212(1)	90	90	90
	93124	ZrO_2	$P4_2/nmc$ S (137)	3.626(2)	3.626(2)	5.235(4)	90	90	90
	93125	ZrO_2	$P4_2/nmc$ S (137)	3.642(1)	3.642(1)	5.275(1)	90	90	90
	93126	ZrO_2	$P4_2/nmc$ S (137)	3.646(1)	3.646(1)	5.285(1)	90	90	90
	94931	ZrO_2	$P4_2/nmc$ Z (137)	3.596	3.596	5.177	90	90	90
	97004	ZrO_2	$P4_2/nmc$ Z (137)	3.598(1)	3.598(1)	5.185(3)	90	90	90
	99744	$Zr_{0.88}O_{1.78}$	$P4_2/nmc$ Z (137)	3.59(6)	3.59(6)	5.16(7)	90	90	90
	157617	ZrO_2	$P4_2/nmc$ S (137)	3.6019(3)	3.6019(3)	5.174(2)	90	90	90
	157618	ZrO_2	$P4_2/nmc$ S (137)	3.6023(6)	3.6023(6)	5.176(4)	90	90	90
	157619	ZrO_2	$P4_2/nmc$ S (137)	3.6005(3)	3.6005(3)	5.177(2)	90	90	90
	157620	ZrO_2	$P4_2/nmc$ S (137)	3.5990(3)	3.5990(3)	5.173(2)	90	90	90
	157621	ZrO_2	$P4_2/nmc$ S (137)	3.6007(3)	3.6007(3)	5.178(2)	90	90	90
	173961	ZrO_2	$P4_2/nmc$ S (137)	3.6287	3.6287	5.207	90	90	90
	180936	ZrO_2	$P4_2/nmc$ Z (137)	3.5975	3.5975	5.1649	90	90	90
	647692	ZrO_2	$P4_2/nmc$ Z (137)	3.6526(15)	3.6526(15)	5.2928(27)	90	90	90
Baddeleyite-ZrO_2(mP12) Srtuc._Num_7	15983	ZrO_2	$P2_1/c$ (14)	5.22	5.27	5.38	90	80.54	90
	18190	ZrO_2	$P2_1/c$ (14)	5.145(5)	5.2075(5)	5.3107(5)	90	99.23(8)	90
	26488	ZrO_2	$P2_1/c$ (14)	5.169(8)	5.232(8)	5.341(8)	90	99.25(17)	90
	41010	ZrO_2	$P2_1/c$ (14)	5.120(22)	5.216(3)	5.281(6)	90	99.01(8)	90
	41572	ZrO_2	$P2_1/c$ (14)	5.1501(2)	5.2077(2)	5.3171(2)	90	99.224(2)	90
	57157	ZrO_2	$P2_1/c$ (14)	5.1463(8)	5.2116(8)	5.3134(8)	90	99.222(1)	90
	57158	ZrO_2	$P2_1/c$ (14)	5.14608(2)	5.21177(2)	5.31301(2)	90	99.222(1)	90
	57451	ZrO_2	$P2_1/c$ (14)	5.1459(4)	5.2115(4)	5.3128(4)	90	99.222(1)	90

	60900	ZrO_2	$P2_1/c$ (14)	5.1507(4)	5.2028(4)	5.3156(4)	90	99.196(4)	90
	60901	ZrO_2	$P2_1/c$ (14)	5.1505(4)	5.2031(4)	5.3154(4)	90	99.194(4)	90
	62993	ZrO_2	$P2_1/c$ (14)	5.1505(1)	5.2116(1)	5.3173(1)	90	99.230(1)	90
	68782	ZrO_2	$P2_1/c$ (14)	5.144(1)	5.133(1)	5.347(1)	90	98.88(2)	90
	72693	$Zr_{0.944}O_2$	$P2_1/c$ (14)	5.14948(19)	5.20211(19)	5.31975(22)	90	99.238(3)	90
	72694	$Zr_{0.932}O_2$	$P2_1/c$ (14)	5.1828(2)	5.2117(2)	5.3731(2)	90	98.835(3)	90
	72695	$Zr_{0.931}O_2$	$P2_1/c$ (14)	5.18549(21)	5.21202(21)	5.37736(20)	90	98.7980(28)	90
	72696	$Zr_{0.936}O_2$	$P2_1/c$ (14)	5.18732(26)	5.21283(26)	5.38047(24)	90	98.7710(36)	90
	72697	$Zr_{0.93}O_2$	$P2_1/c$ (14)	5.18674(43)	5.21267(46)	5.38074(44)	90	98.7540(68)	90
	72698	$Zr_{0.93}O_2$	$P2_1/c$ (14)	5.18733(91)	5.2146(10)	5.38402(90)	90	98.729(14)	90
	80042	ZrO_2	$P2_1/c$ (14)	5.1482(2)	5.2076(2)	5.3149(2)	90	99.232(2)	90
	80043	ZrO_2	$P2_1/c$ (14)	5.1484(2)	5.2067(2)	5.3154(2)	90	99.229(2)	90
	80044	ZrO_2	$P2_1/c$ (14)	5.1495(40)	5.2075(40)	5.3164(40)	90	99.23(6)	90
	80045	ZrO_2	$P2_1/c$ (14)	5.1496(1)	5.2076(2)	5.3163(2)	90	99.225(1)	90
	80046	ZrO_2	$P2_1/c$ (14)	5.1491(3)	5.2079(3)	5.3160(3)	90	99.225(2)	90
	80047	ZrO_2	$P2_1/c$ (14)	5.1505(2)	5.2077(3)	5.3164(3)	90	99.223(2)	90
	80048	ZrO_2	$P2_1/c$ (14)	5.1481(2)	5.2059(2)	5.3209(2)	90	99.216(1)	90
	80049	ZrO_2	$P2_1/c$ (14)	5.149(1)	5.2026(9)	5.313(1)	90	99.23(2)	90
	80050	ZrO_2	$P2_1/c$ (14)	5.1423(9)	5.2000(1)	5.311(1)	90	99.205(6)	90
	82543	ZrO_2	$P2_1/c$ (14)	5.14422(4)	5.20969(5)	5.31120(50)	90	99.220(1)	90
	82544	ZrO_2	$P2_1/c$ (14)	5.14604(6)	5.21162(7)	5.31308(70)	90	99.222(1)	90
	82545	ZrO_2	$P2_1/c$ (14)	5.14478(4)	5.21028(4)	5.31179(40)	90	99.226(6)	90
	86692	ZrO_2	$P2_1/c$ (14)	5.1489(1)	5.2118(1)	5.3147(1)	90	99.203(2)	90
	89426	ZrO_2	$P2_1/c$ (14)	5.146(7)	5.205(1)	5.313(6)	90	99.1(8)	90
	94886	ZrO_2	$P2_1/c$ (14)	5.1451(3)	5.2023(4)	5.3219(4)	90	99.15(3)	90
	94887	ZrO_2	$P2_1/c$ (14)	5.1487(4)	5.2023(4)	5.3231(4)	90	99.164(6)	90
	96537	ZrO_2	$P2_1/c$ (14)	5.1510(4)	5.2031(4)	5.3151(4)	90	99.197(2)	90
	157403	ZrO_2	$P2_1/c$ (14)	5.1514(4)	5.2098(4)	5.3204(4)	90	99.171(3)	90
	172161	ZrO_2	$P2_1/c$ (14)	5.1462(20)	5.2082(23)	5.3155(24)	90	99.249(35)	90
	173959	ZrO_2	$P2_1/c$ (14)	5.1974	5.298	5.3498	90	99.53	90
	291120	ZrO_2	$P2_1/c$ (14)	5.195(9)	5.225(4)	5.342(8)	90	98.17	90
	417639	ZrO_2	$P2_1/c$ (14)	5.14281(9)	5.20368(4)	5.30971(10)	90	99.166(6)	90
	647691	ZrO_2	$P2_1/c$ (14)	5.1415(5)	5.2056(5)	5.3128(5)	90	99.30(8)	90
	658755	ZrO_2	$P2_1/c$ (14)	5.1464(5)	5.2116(5)	5.3133(5)	90	99.221(1)	90
	659226	ZrO_2	$P2_1/c$ (14)	5.09	5.13	5.29	90	99.89	90
Fluorite-CaF_2, Fluorite-CaF_2(defect), NaCl, ZrO Srtuc._Num_8	28991	$La_{0.5}Zr_{0.5}O_{1.75}$	Fm-3m (225)	5.407	5.407	5.407	90	90	90
	53998	ZrO_2	Fm-3m (225)	5.1	5.1	5.1	90	90	90
	72955	$ZrO_{2.12}$	Fm-3m (225)	5.1291(8)	5.1291(8)	5.1291(8)	90	90	90
	72956	$ZrO_{1.87}$	Fm-3m (225)	5.1523(7)	5.1523(7)	5.1523(7)	90	90	90
	75064	$(La_{0.1}Zr_{0.9})O_{1.95}$	Fm-3m (225)	5.1149(20)	5.1149(20)	5.1149(20)	90	90	90
	76019	ZrO	Fm-3m (225)	4.602	4.602	4.602	90	90	90
	77713	ZrO	Fm-3m (225)	5.11(4)	5.11	5.11	90	90	90
	81035	$(Zr_{0.9}La_{0.1})O_{1.95}$	Fm-3m (225)	5.1635(3)	5.1635(3)	5.1635(3)	90	90	90
	89429	ZrO_2	Fm-3m (225)	5.135(9)	5.135(9)	5.135(9)	90	90	90
	92095	ZrO_2	Fm-3m (225)	4.925(2)	4.925(2)	4.925(2)	90	90	90
	92096	ZrO_2	Fm-3m (225)	4.916(2)	4.916(2)	4.916(2)	90	90	90
	105553	ZrO_2	Fm-3m (225)	5.09	5.09	5.09	90	90	90
	173962	ZrO_2	Fm-3m (225)	5.128	5.128	5.128	90	90	90
	253062	$(La_{0.5}Zr_{0.5})O_{1.75}$	Fm-3m (225)	5.385(2)	5.385(2)	5.385(2)	90	90	90
	291119	ZrO_2	Fm-3m (225)	5.145(8)	5.145(8)	5.145(8)	90	90	90
	647689	ZrO_2	Fm-3m (225)	5.075(10)	5.075(10)	5.075(10)	90	90	90
ZrO_2(oP12) Srtuc._Num_9	41011	ZrO_2	Pbcm (57)	5.005(8)	5.235(2)	5.051(2)	90	90	90
	41012	ZrO_2	Pbcm (57)	4.992(7)	5.229(2)	5.046(2)	90	90	90

	67004	ZrO_2	$Pbc2_1$ (29)	5.068(1)	5.260(1)	5.077(1)	90	90	90
	77716	ZrO_2	Pbcm (57)	5.0364(1)	5.2546(1)	5.0855(1)	90	90	90
HfO_2 Srtuc._Num_10	56696	ZrO_2	Pnma (62)	5.520(4)	3.47(4)	6.503(4)	90	90	90
	79914	ZrO_2	Pnam (62)	5.593(1)	6.484(1)	3.333(1)	90	90	90
	79915	ZrO_2	Pnam (62)	5.471(3)	6.341(4)	3.246(1)	90	90	90
	83862	ZrO_2	Pnam (62)	5.5873(2)	6.4847(2)	3.3298(1)	90	90	90
	173963	ZrO_2	Pnma (62)	5.614	3.3474	6.5658	90	90	90
Brookite-TiO_2 Srtuc._Num_11	173960	ZrO_2	Pbca (61)	10.1745	5.3148	5.1357	90	90	90
$SmLiO_2$ Srtuc._Num_12	239278	$LaLiO_2$	$P2_1/c$ (14)	5.874(4)	6.2129(4)	5.8481(7)	90	102.453(3)	90
No structure type (Li_La_O) Srtuc._Num_13	259020	$La_2LiO_{3.5}$	Fmmm (69)	5.196(6)	5.136(7)	13.279(7)	90	90	90
Pyrochlore-$NaCa(Nb_2O_6)F$, $Eu_2Zr_2O_7$ Srtuc._Num_14	15165	$La_2(Zr_2O_7)$	Fd-3mS (227)	10.786	10.786	10.786	90	90	90
	51573	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.7997(3)	10.7997(3)	10.7997(3)	90	90	90
	150206	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.802(1)	10.802(1)	10.802(1)	90	90	90
	153222	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.80470(6)	10.80470(6)	10.80470(6)	90	90	90
	173795	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.8076(5)	10.8076(5)	10.8076(5)	90	90	90
	180491	$La_2(Zr_2O_7)$	Fd-3mS (227)	10.808(1)	10.808(1)	10.808(1)	90	90	90
	184089	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.811(4)	10.811(4)	10.811(4)	90	90	90
	184090	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.828(4)	10.828(4)	10.828(4)	90	90	90
	184091	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.848(4)	10.848(4)	10.848(4)	90	90	90
	184092	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.869(4)	10.869(4)	10.869(4)	90	90	90
	184093	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.892(3)	10.892(3)	10.892(3)	90	90	90
	184094	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.910(4)	10.910(4)	10.910(4)	90	90	90
	184095	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.915(4)	10.915(4)	10.915(4)	90	90	90
	184096	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.921(4)	10.921(4)	10.921(4)	90	90	90
	184097	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.927(4)	10.927(4)	10.927(4)	90	90	90
	184098	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.933(4)	10.933(4)	10.933(4)	90	90	90
	184099	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.939(3)	10.939(3)	10.939(3)	90	90	90
	184100	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.945(4)	10.945(4)	10.945(4)	90	90	90
	184101	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.951(4)	10.951(4)	10.951(4)	90	90	90
	184102	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.957(4)	10.957(4)	10.957(4)	90	90	90
	189341	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.812(1)	10.812(1)	10.812(1)	90	90	90
	189344	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.809(1)	10.809(1)	10.809(1)	90	90	90
	248648	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.875(1)	10.875(1)	10.875(1)	90	90	90
	248649	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.906(1)	10.906(1)	10.906(1)	90	90	90
	248878	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.78864(10)	10.78864(10)	10.78864(10)	90	90	90
	253061	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.799(8)	10.799(8)	10.799(8)	90	90	90
	253063	$La_2(Zr_2O_7)$	Fd-3mZ (227)	10.798(3)	10.798(3)	10.798(3)	90	90	90
Li_2PbO_3, Cu_2GeS_3 Srtuc._Num_15	31941	Li_2ZrO_3	Cc (9)	5.427	9.025	5.427	90	112.75	90
	35236	$Li_{1.82}ZrO_3$	C2/c (15)	5.4218(2)	9.0216(4)	5.4187(2)	90	112.709(2)	90
	94565	Li_2ZrO_3	C2/c (15)	5.5208(5)	9.0759(11)	5.4758(5)	90	113.485(7)	90
	94893	Li_2ZrO_3	C2/c (15)	5.4089(1)	9.0309(2)	5.4144(1)	90	112.498(1)	90
	94894	Li_2ZrO_3	C2/c (15)	5.4163(1)	9.0262(2)	5.4156(1)	90	112.615(1)	90
	94895	Li_2ZrO_3	C2/c (15)	5.4240(1)	9.0263(2)	5.4197(1)	90	112.701(1)	90
	94896	Li_2ZrO_3	C2/c (15)	5.4392(1)	9.0383(2)	5.4302(1)	90	112.791(1)	90
	94897	Li_2ZrO_3	C2/c (15)	5.4686(1)	9.0555(1)	5.4477(1)	90	113.002(1)	90
	241300	$Li_{1.85}ZrO_3$	C2/c (15)	5.4268(1)	9.0201(3)	5.4184(6)	90	112.774(3)	90
	241301	$Li_{1.8}ZrO_3$	C2/c (15)	5.4253(3)	9.0229(4)	5.4186(6)	90	112.737(2)	90
	291154	Li_2ZrO_3	C2/c (15)	5.4264(1)	9.0258(2)	5.4211(1)	90	112.7234(8)	90
$Li_6Zr_2O_7$ Srtuc._Num_16	41321	$Li_6Zr_2O_7$	C2/c (15)	10.445(1)	5.989(1)	10.200(1)	90	100.26(1)	90
	73835	$Li_6Zr_2O_7$	C2/c (15)	10.440(4)	5.991(1)	10.204(2)	90	100.25(3)	90
	73836	$Li_6Zr_2O_7$	C2/c (15)	10.4428(1)	5.9877(1)	10.2014(1)	90	100.266(1)	90
Li_8IrO_6	61217	$Li_8(IrO_6)$	R-3H (148)	5.4151(6)	5.4151(6)	15.0584(37)	90	90	120

Srtuc._Num_17									
No structure type (La_Zr_O) Srtuc._Num_18	203085	$Li_4Zr_3O_8$	$I4_1$/amdZ (141)	4.265	4.265	9.0149	90	90	90
	183684	$Li_7La_3Zr_2O_{12}$	$I4_1$/acdZ (142)	13.1189(4)	13.1189(4)	12.6701(4)	90	90	90
	183685	$Li_7La_3Zr_2O_{12}$	$I4_1$/acdZ (142)	13.0920(9)	13.0920(9)	12.618(1)	90	90	90
	191528	$Li_7La_3Zr_2O_{12}$	$I4_1$/acdZ (142)	13.1065(1)	13.1065(1)	12.6143(1)	90	90	90
$Li_7La_3Sn_2O_{12}$ Srtuc._Num_19	238686	$Li_7La_3Zr_2O_{12}$	$I4_1$/acdZ (142)	13.1954(2)	13.1954(2)	12.8686(3)	90	90	90
	238687	$Li_7La_3Zr_2O_{12}$	$I4_1$/acdZ (142)	13.1225(1)	13.1225(1)	12.6674(2)	90	90	90
	246816	$Li_7La_3Zr_2O_{12}$	$I4_1$/acdZ (142)	13.134(4)	13.134(4)	12.663(8)	90	90	90
	246817	$Li_7La_3Zr_2O_{12}$	$I4_1$/acdZ (142)	13.1279(5)	13.1279(5)	12.6715(5)	90	90	90
	183607	$Li_7La_3Zr_2O_{12}$	Ia-3d (230)	13.0035	13.0035	13.0035	90	90	90
	184230	$Li_7La_3Zr_2O_{12}$	Ia-3d (230)	13.0035	13.0035	13.0035	90	90	90
$Li_7La_3Zr_2O_{12}$ Srtuc._Num_20	238685	$Li_7La_3Zr_2O_{12}$	Ia-3d (230)	13.1842(3)	13.1842(3)	13.1842(3)	90	90	90
	261302	$Li_7La_3Zr_2O_{12}$	Ia-3d (230)	12.9751(1)	12.9751(1)	12.9751(1)	90	90	90
	422259	$Li_7La_3Zr_2O_{12}$	Ia-3d (230)	12.9827(4)	12.9827(4)	12.9827(4)	90	90	90

Table S1 The list of 218 constituent compounds. The structure type and the number of duplicates for each of 21 groups (**Group_Num_0~20**) are also given.

Phase identification accuracy comparison

	Li-La-Zr-O	Li-Sr-Al-O
Average Pearson correlation coefficient	0.4899	0.5130
CNN classification	96.47% (91.11%)	99.30% (98.67%)
KNN	13.08% (24.44%)	35.72% (8.00%)
RF	63.62% (17.78%)	87.63% (58.00%)
SVM	42.74% (13.33%)	45.76% (18.00%)

Table S2 The phase identification accuracy comparison between the present (Li-La-Zr-O composition) and previous (Li-Sr-Al-O composition) approaches. The numbers in the parentheses designate the real-data test accuracy. The values presented here indicate the best accuracy obtained from each of the algorithms. The hyper-parameters used for the KNN, RF, and SVM algorithms are the same as those for the phase fraction regression described in the caption of Figure 2.

Ground-Truth Phase Fraction Data

Sample Number	Compound			Compound Fraction		
1	La_2O_3	$LaLiO_2$	X	0.8671	0.1329	0
2	La_2O_3	$La_2Zr_2O_7$	X	0.8006	0.1994	0
3	La_2O_3	$LaLiO_2$	X	0.7973	0.2027	0
4	La_2O_3	$La_2Zr_2O_7$	X	0.8222	0.1778	0
5	La_2O_3	$La_2Zr_2O_7$	X	0.6125	0.3875	0
6	La_2O_3	ZrO_2	X	0.8945	0.1055	0
7	La_2O_3	$La_2Zr_2O_7$	X	0.4908	0.5092	0
8	La_2O_3	$La_2Zr_2O_7$	X	0.546	0.454	0
9	La_2O_3	ZrO_2	X	0.8547	0.1453	0
10	La_2O_3	$Li_7La_3Zr_2O_{12}$	X	0.8555	0.1445	0
11	La_2O_3	$La_2Zr_2O_7$	$Li_7La_3Zr_2O_{12}$	0.4709	0.4343	0.0948
12	La_2O_3	$La_2Zr_2O_7$	X	0.4537	0.5463	0
13	La_2O_3	$La_2Zr_2O_7$	X	0.4055	0.5945	0
14	La_2O_3	ZrO_2	$La_2Zr_2O_7$	0.464	0.1575	0.3785
15	La_2O_3	$LaLiO_2$	$Li_7La_3Zr_2O_{12}$	0.5485	0.1583	0.2932
16	La_2O_3	$La_2Zr_2O_7$	X	0.3295	0.6705	0
17	La_2O_3	$La_2Zr_2O_7$	X	0.2196	0.7804	0
18	Li_2O	La_2O_3	$La_2Zr_2O_7$	0.1891	0.2068	0.604
19	La_2O_3	ZrO_2	$La_2Zr_2O_7$	0.1783	0.1192	0.7026
20	La_2O_3	$La_2Zr_2O_7$	$Li_7La_3Zr_2O_{12}$	0.2604	0.5771	0.1625
21	La_2O_3	$La_2Zr_2O_7$	X	0.2667	0.7333	0
22	La_2O_3	ZrO_2	X	0.6129	0.3871	0
23	$La_2Zr_2O_7$	Li_2ZrO_3	X	0.8334	0.1666	0
24	La_2O_3	$La_2Zr_2O_7$	Li_2ZrO_3	0.2103	0.6535	0.1361
25	La_2O_3	ZrO_2	$La_2Zr_2O_7$	0.3773	0.5182	0.1045
26	Li_2O	Li_8ZrO_6	$Li_7La_3Zr_2O_{12}$	0.1342	0.3455	0.5203
27	La_2O_3	$La_2Zr_2O_7$	Li_2ZrO_3	0.3896	0.1163	0.4941
28	ZrO_2	$La_2Zr_2O_7$	Li_2ZrO_3	0.5029	0.3774	0.1197
29	ZrO_2	$La_2Zr_2O_7$	X	0.6226	0.3774	0
30	ZrO_2	$La_2Zr_2O_7$	X	0.6647	0.3353	0
31	La_2O_3	ZrO_2	X	0.2426	0.7574	0
32	$Li_6Zr_2O_7$	Li_8ZrO_6	X	0.5901	0.4099	0
33	ZrO_2	Li_2ZrO_3	X	0.2928	0.7072	0
34	ZrO_2	Li_2ZrO_3	X	0.5707	0.4293	0
35	ZrO_2	Li_2ZrO_3	X	0.8289	0.1711	0
36	ZrO_2	X	X	0.9041	0.0959	0
37	La_2O_3	X	X	1	0	0

38	La_2O_3	X	X	1	0	0
39	La_2O_3	X	X	1	0	0
40	La_2O_3	X	X	1	0	0
41	La_2O_3	X	X	1	0	0
42	Li_2O	X	X	1	0	0
43	Li_8ZrO_6	X	X	1	0	0
44	ZrO_2	X	X	1	0	0
45	ZrO_2	X	X	1	0	0

Table S3 The phase identification and the relative phase fraction evaluation results for 45 real-world XRD patterns. The conventional rule-based XRD analysis tools such as X'pert pro and FullProf were used to produce the following ground-truth data.

Real-world data phase identification test accuracy

Dataset	D1			D2		
CNN architecture	**CNN_4**	**CNN_5**	**CNN_6**	**CNN_4**	**CNN_5**	**CNN_6**
10th epoch	86.67% (95.56%)	80.00% (93.33%)	88.88% (96.30%)	77.78% (92.59%)	80.00% (93.33%)	91.11% (97.03%)
16th epoch	86.67% (95.56%)	82.22% (94.07%)	86.67% (95.56%)	75.56% (91.85%)	77.78% (92.59%)	80.00% (93.33%)
17th epoch	86.67% (95.56%)	86.67% (95.56%)	86.67% (95.56%)	82.22% (94.07%)	73.33% (91.11%)	82.22% (94.07%)
18th epoch	77.78% (92.59%)	84.44% (94.81%)	86.67% (95.56%)	77.78% (92.59%)	80.00% (93.33%)	86.67% (95.56%)
19th epoch	84.44% (94.81%)	84.44% (94.81%)	86.67% (95.56%)	80.00% (93.33%)	71.11% (90.37%)	80.00% (93.33%)
20th epoch	80.00% (93.33%)	82.22% (94.07%)	84.44% (94.81%)	73.33% (91.11%)	71.11% (90.37%)	80.00% (93.33%)

Table S4 Real-world data phase identification test accuracy for CNN_n (n=4~6) trained with D1 and D2 datasets, using the weight and bias saved at the 10^{th} and 16^{th}~20^{th} epochs. The total constituent population-based accuracy values are given in the parentheses.

Synthetic data phase identification test accuracy

Dataset	D1			D2		
CNN architecture	CNN_4	CNN_5	CNN_6	CNN_4	CNN_5	CNN_6
10^{th} epoch	91.05% (97.00%)	91.88% (97.27%)	90.76% (96.90%)	94.56% (98.18%)	93.74% (97.90%)	93.82% (97.93%)
16^{th} epoch	92.67% (97.54%)	93.63% (97.86%)	92.36% (97.44%)	96.06% (98.68%)	94.64% (98.21%)	94.84% (98.28%)
17^{th} epoch	93.18% (97.72%)	93.74% (97.90%)	92.45% (97.47%)	96.10% (98.70%)	95.01% (98.33%)	95.18% (98.39%)
18^{th} epoch	93.30% (97.76%)	93.83% (97.94%)	92.76% (97.57%)	96.33% (98.77%)	95.10% (98.36%)	95.46% (98.48%)
19^{th} epoch	93.15% (97.71%)	94.19% (98.06%)	93.02% (97.66%)	96.07% (98.69%)	95.24% (98.41%)	95.61% (98.53%)
20^{th} epoch	93.42% (97.80%)	94.36% (98.11%)	93.04% (97.67%)	96.47% (98.82%)	95.70% (98.56%)	95.42% (98.47%)

Table S5 Synthetic data phase identification test accuracy for CNN_n (n=4~6) trained with D1 and D2 datasets, using the weight and bias saved at the 10^{th} and 16^{th}~20^{th} epochs. The total constituent population-based accuracy values are given in the parentheses.