Electronic supplemental information for

"Predictive nonlinear optical crystal formation energy regression model based on convolutional neural networks"

Zhen Fan,[*a] , Shuai Lian[b], Guangyong Jin,[a] Chao Xin[*a], Ye Li[a], and Boshi Yuan,[*a]

a. School of Science, Changchun University of Science and Technology, Changchun 130022, China.

b. School of Medical Technology, Beijing Institute of Technology, Beijing 100080, China

Corresponding Author

fz@cust.edu.cn (Zhen Fan)

xinchao@cust.edu.cn (Chao Xin)

yuanboshi1986@163.com (Boshi Yuan)

Table S1 Results of hyperparameter testing

Hyperparameters	Train_R2	Train_RMSE	Train_MAE	Test_R2	Test_RMSE	Test_MAE
batchsize8_outchannel7_neure1st40_neure2sec10	0.999	0.034	0.023	0.973	0.171	0.103
batchsize8_outchannel8_neure1st40_neure2sec10	0.997	0.057	0.042	0.977	0.156	0.104
batchsize8_outchannel9_neure1st40_neure2sec10	0.999	0.038	0.027	0.981	0.143	0.084
batchsize8_outchannel7_neure1st40_neure2sec20	0.997	0.055	0.039	0.981	0.143	0.097
batchsize8_outchannel8_neure1st40_neure2sec20	0.998	0.045	0.032	0.98	0.148	0.092
batchsize8_outchannel9_neure1st40_neure2sec20	0.998	0.042	0.031	0.978	0.153	0.095
batchsize8_outchannel7_neure1st40_neure2sec30	0.997	0.062	0.045	0.978	0.152	0.106
batchsize8_outchannel8_neure1st40_neure2sec30	0.998	0.046	0.03	0.984	0.133	0.088
batchsize8_outchannel9_neure1st40_neure2sec30	0.999	0.04	0.03	0.98	0.148	0.092
batchsize8_outchannel7_neure1st50_neure2sec10	0.997	0.055	0.039	0.981	0.143	0.096
batchsize8_outchannel8_neure1st50_neure2sec10	0.999	0.031	0.022	0.981	0.143	0.091
batchsize8_outchannel9_neure1st50_neure2sec10	0.998	0.046	0.032	0.983	0.136	0.092
batchsize8_outchannel7_neure1st50_neure2sec20	0.999	0.038	0.026	0.976	0.162	0.103
batchsize8_outchannel8_neure1st50_neure2sec20	0.999	0.038	0.029	0.979	0.151	0.094
batchsize8_outchannel9_neure1st50_neure2sec20	0.999	0.041	0.028	0.983	0.135	0.088
batchsize8_outchannel7_neure1st50_neure2sec30	0.998	0.051	0.036	0.981	0.144	0.091
batchsize8_outchannel8_neure1st50_neure2sec30	0.999	0.033	0.025	0.982	0.141	0.092
batchsize8_outchannel9_neure1st50_neure2sec30	0.998	0.042	0.031	0.979	0.15	0.093
batchsize8_outchannel7_neure1st60_neure2sec10	0.999	0.034	0.024	0.977	0.156	0.096
batchsize8_outchannel8_neure1st60_ne	0.997	0.058	0.042	0.983	0.136	0.092

ure2sec10						
batchsize8_outchannel9_neure1st60_neure2sec10	0.998	0.042	0.03	0.982	0.139	0.088
batchsize8_outchannel7_neure1st60_neure2sec20	0.999	0.034	0.024	0.979	0.15	0.094
batchsize8_outchannel8_neure1st60_neure2sec20	0.999	0.038	0.026	0.98	0.146	0.093
batchsize8_outchannel9_neure1st60_neure2sec20	0.999	0.033	0.023	0.98	0.148	0.09
batchsize8_outchannel7_neure1st60_neure2sec30	0.998	0.042	0.031	0.979	0.152	0.096
batchsize8_outchannel8_neure1st60_neure2sec30	0.998	0.047	0.035	0.979	0.15	0.101
batchsize8_outchannel9_neure1st60_neure2sec30	0.999	0.029	0.019	0.984	0.131	0.085
batchsize16_outchannel7_neure1st40_neure2sec10	0.999	0.035	0.024	0.982	0.138	0.087
batchsize16_outchannel8_neure1st40_neure2sec10	-0.102	1.114	0.889	-0.086	1.082	0.872
batchsize16_outchannel9_neure1st40_neure2sec10	0.999	0.034	0.026	0.976	0.161	0.096
batchsize16_outchannel7_neure1st40_neure2sec20	0.998	0.044	0.035	0.981	0.144	0.093
batchsize16_outchannel8_neure1st40_neure2sec20	0.998	0.052	0.036	0.976	0.16	0.104
batchsize16_outchannel9_neure1st40_neure2sec20	0.999	0.037	0.028	0.982	0.139	0.093
batchsize16_outchannel7_neure1st40_neure2sec30	0.998	0.041	0.03	0.972	0.174	0.103
batchsize16_outchannel8_neure1st40_neure2sec30	0.998	0.042	0.029	0.98	0.147	0.093
batchsize16_outchannel9_neure1st40_neure2sec30	0.998	0.048	0.034	0.981	0.142	0.095
batchsize16_outchannel7_neure1st50_neure2sec10	0.999	0.038	0.027	0.982	0.14	0.094
batchsize16_outchannel8_neure1st50_neure2sec10	0.998	0.052	0.038	0.981	0.142	0.098
batchsize16_outchannel9_neure1st50_neure2sec10	0.999	0.034	0.024	0.983	0.135	0.086
batchsize16_outchannel7_neure1st50_neure2sec20	0.999	0.035	0.024	0.982	0.14	0.089
batchsize16_outchannel8_neure1st50_neure2sec20	0.999	0.035	0.025	0.984	0.133	0.087

batchsize16_outchannel9_neure1st50_neure2sec20	0.998	0.045	0.032	0.984	0.133	0.087
batchsize16_outchannel7_neure1st50_neure2sec30	0.997	0.054	0.04	0.982	0.141	0.094
batchsize16_outchannel8_neure1st50_neure2sec30	0.999	0.033	0.023	0.982	0.138	0.087
batchsize16_outchannel9_neure1st50_neure2sec30	0.999	0.033	0.023	0.984	0.131	0.085
batchsize16_outchannel7_neure1st60_neure2sec10	0.998	0.046	0.033	0.977	0.156	0.094
batchsize16_outchannel8_neure1st60_neure2sec10	0.998	0.042	0.029	0.982	0.139	0.093
batchsize16_outchannel9_neure1st60_neure2sec10	0.999	0.033	0.024	0.984	0.131	0.084
batchsize16_outchannel7_neure1st60_neure2sec20	0.999	0.034	0.024	0.985	0.128	0.083
batchsize16_outchannel8_neure1st60_neure2sec20	0.997	0.062	0.044	0.98	0.146	0.1
batchsize16_outchannel9_neure1st60_neure2sec20	0.997	0.057	0.04	0.979	0.15	0.097
batchsize16_outchannel7_neure1st60_neure2sec30	0.998	0.043	0.032	0.982	0.139	0.093
batchsize16_outchannel8_neure1st60_neure2sec30	0.999	0.031	0.022	0.983	0.134	0.087
batchsize16_outchannel9_neure1st60_neure2sec30	0.999	0.038	0.027	0.982	0.141	0.091

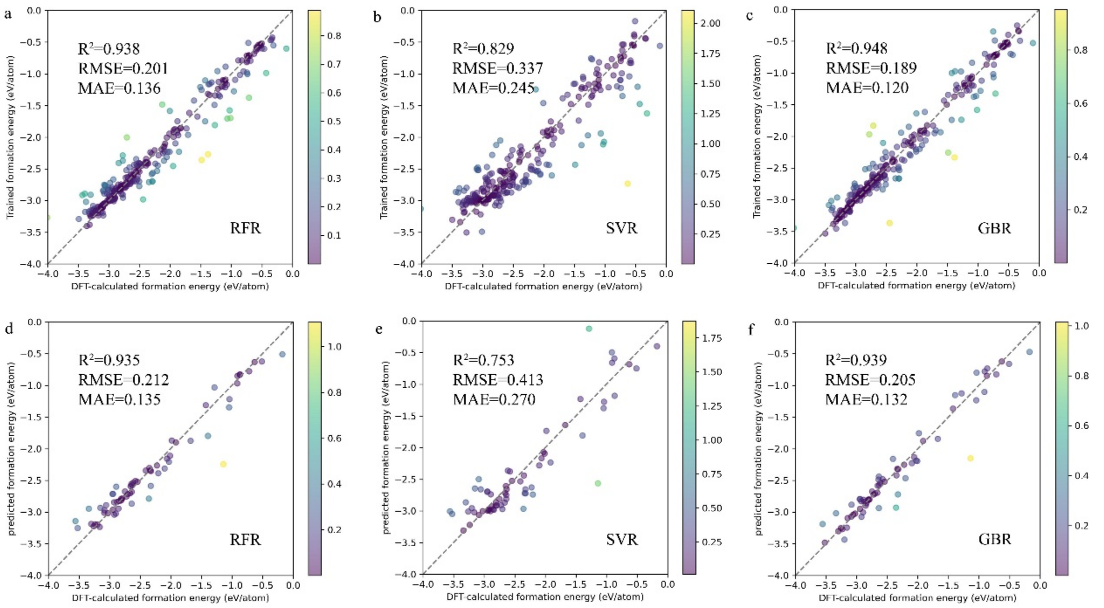

Figure S1 Our previous work, we employed Random Forest Regression (RFR), Support Vector Machine Regression (SVR), and Gradient Boosting Regression (GBR) to forecast the formation energy of nonlinear optical crystals. The training results of RFR (a), SVR(b), and GBR(c) models, and testing results of RFR (d), SVR(e), and GBR(f) models. The gray dash line in each figure represents the ideal curve y = x. The color bar represents absolute error.

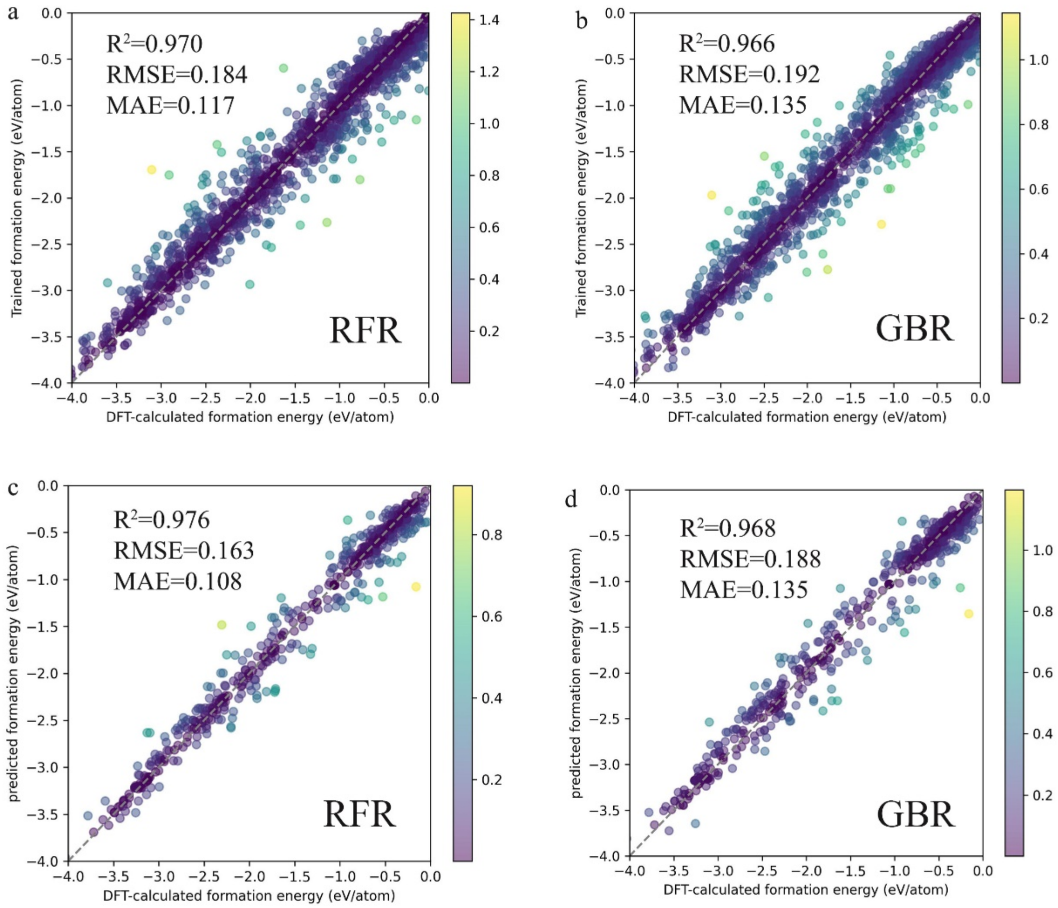

Figure S2 The crystal dataset developed in this paper is employed to apply random forest regression (RFR) and gradient boosting regression (GBR) techniques for forecasting the formation energy of nonlinear optical crystals. The training results of RFR (a), and GBR(b) models, and testing results of RFR (c) and GBR(d) models. The gray dash line in each figure represents the ideal curve y = x. The color bar represents absolute error.