GrapheNet: a deep learning framework for predicting the physical and electronic properties of nanographenes using images

Tommaso Forni[1,2], Matteo Baldoni[1], Fabio Le Piane[1,3] and Francesco Mercuri[1*]

[1]DAIMON Lab, Istituto per lo Studio dei Materiali Nanostrutturati (ISMN), Consiglio Nazionale delle Ricerche (CNR), Bologna, 40129, Italy
[2]Department of Control and Computer Engineering, Polytechnic University of Turin, Corso Castelfidardo 34/d, Turin, 10138, Italy
[3]Department of Computer Science and Engineering, University of Bologna, via Zamboni 33, Bologna, 40126, Italy

*Corresponding author. Email: francesco.mercuri@cnr.it

Supporting Information

Supplementary figures

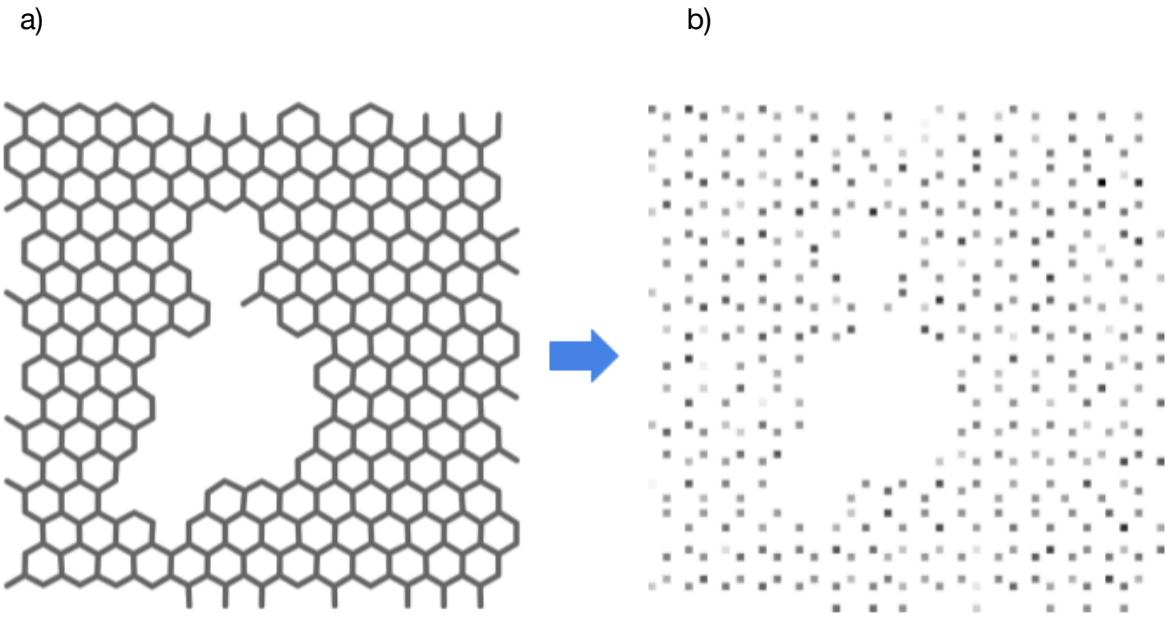


Figure S1. Mapping of samples from the DG dataset. a) Structural representation of a defected graphene flake sample extracted from the DG dataset, rendered with VMD[1]; b) The same sample as it appears in the GrapheNet representation as a grayscale image, generated with the proposed approach (see Methods) and plotted with OpenCV[2] and Pillow[3] (in the case of the DG dataset there is only one atom type).

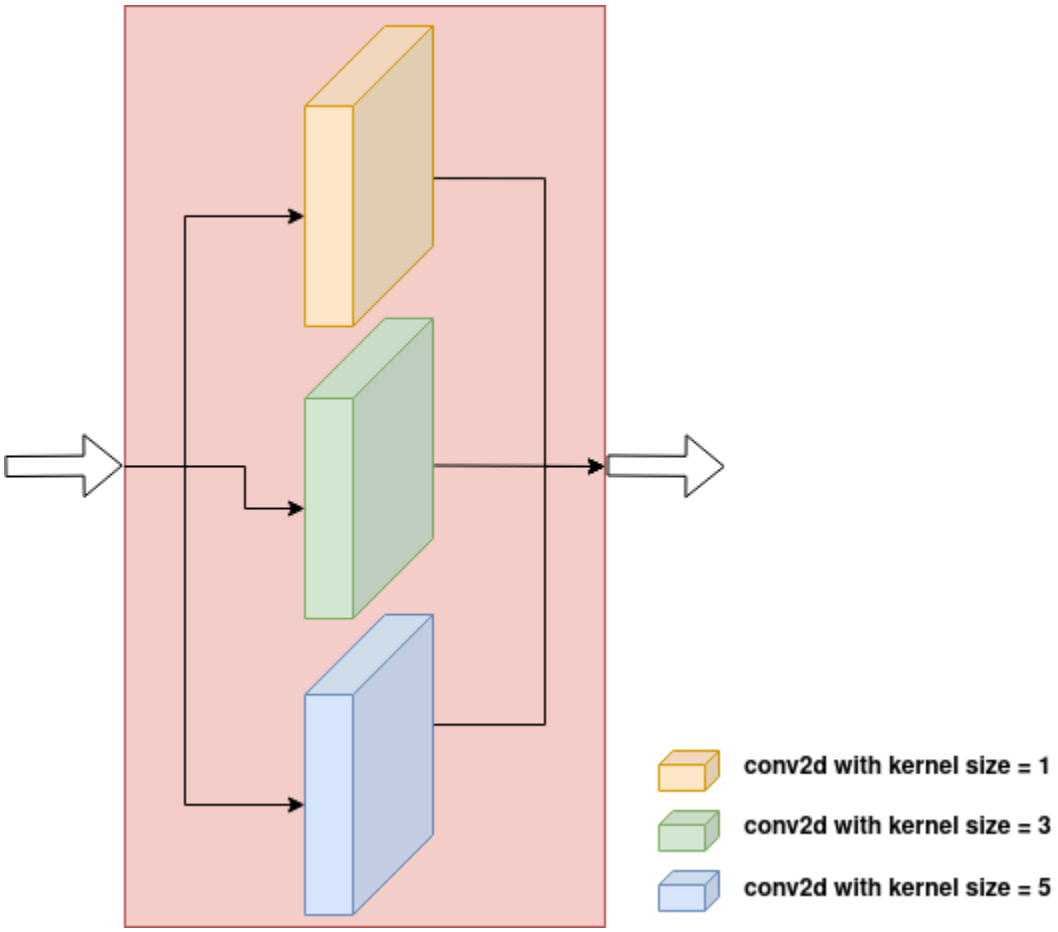


Figure S2. Architecture of an InceptionBlock, plotted with drawio[4], designed to capture information at different spatial scales by using multiple convolutional filters of different sizes in parallel. In our work, we considered an Inception block with three types of convolutional filters: 1x1, 3x3, and 5x5.

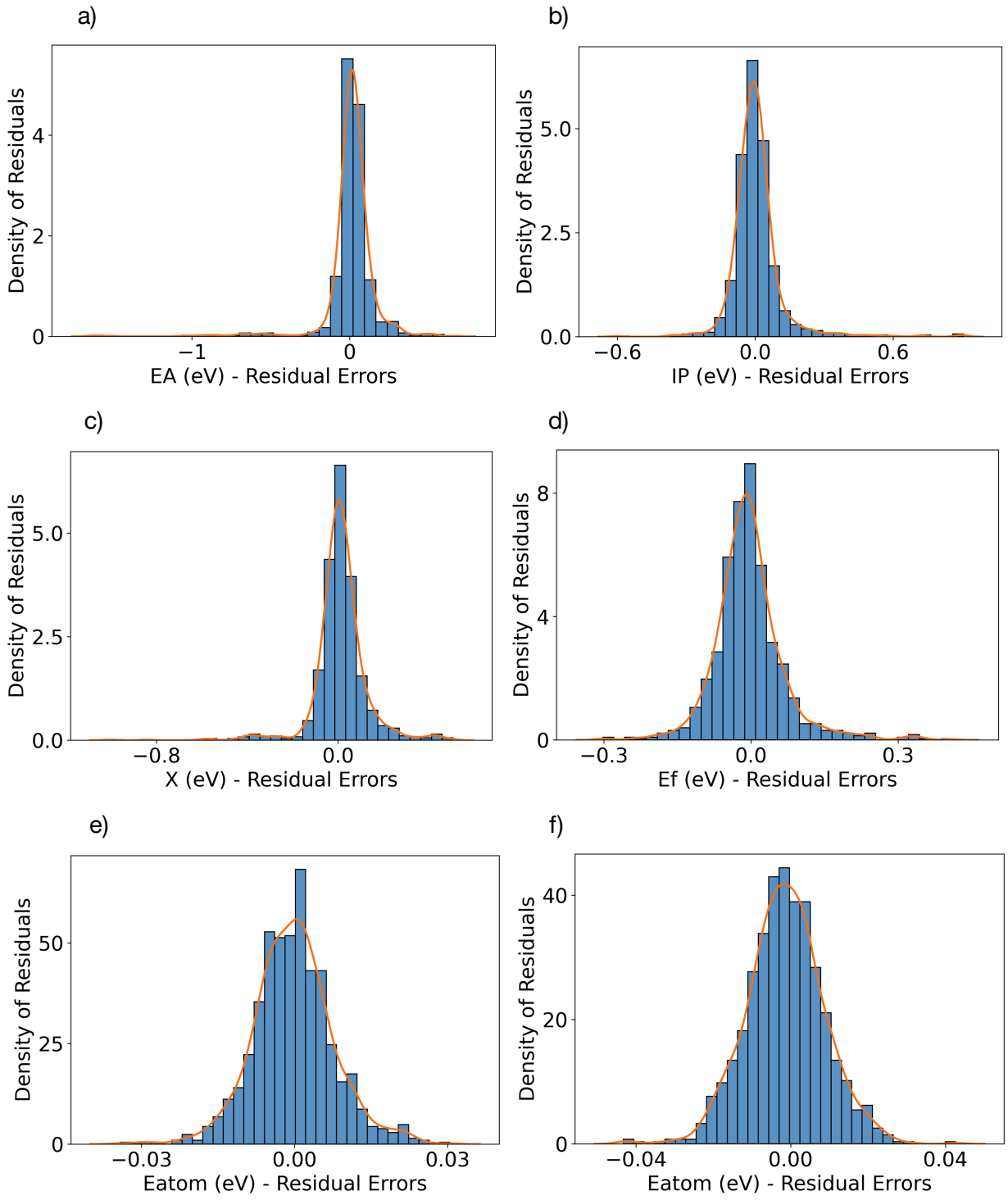


Figure S3. Residual errors for the predictions of: a) electron affinity, b) ionization potential, c) electronegativity, d) Fermi energy and e) formation energy per atom for the GO dataset and f) formation energy per atom for the DG dataset.

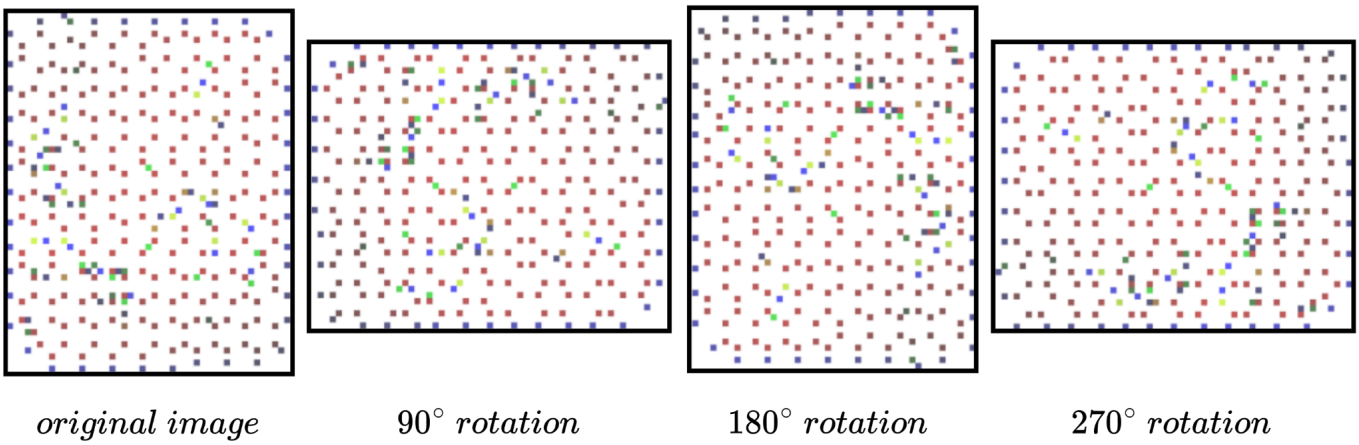


Figure S4. Data augmentation: from the original image, 3 additional images are obtained by applying three-fold rotation operations by multiples of 90°, in order to have a total of 4 images representation for each sample of the dataset. The rotation operation has been performed using OpenCV[2].

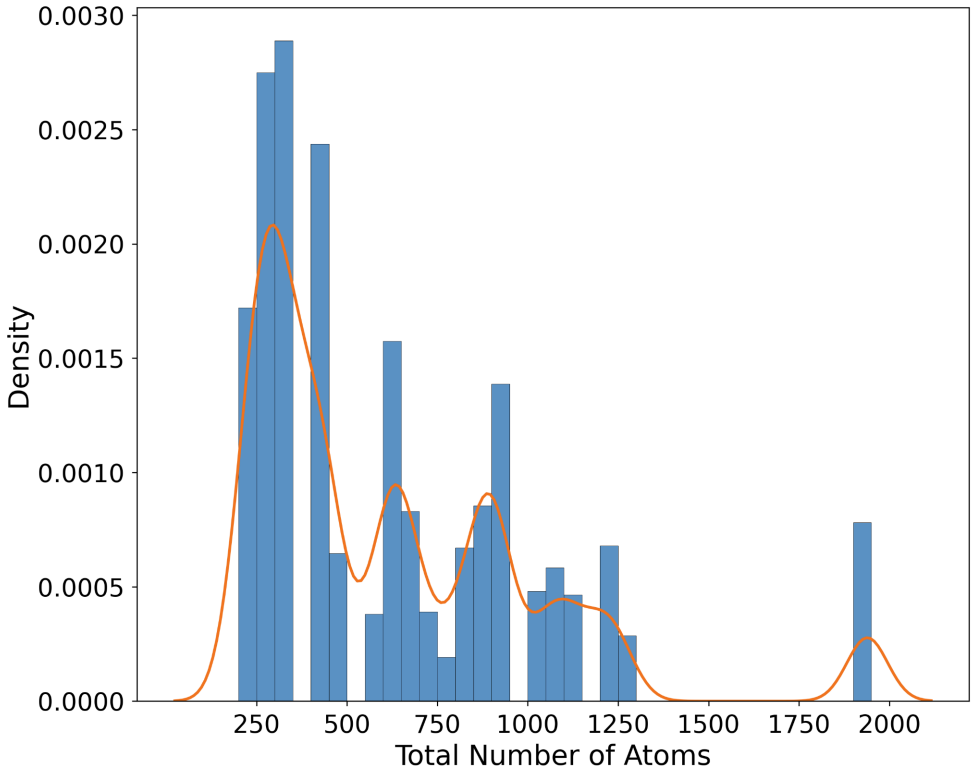


Figure S5. Distribution of the total number of atoms per flakes of the complete GO dataset.

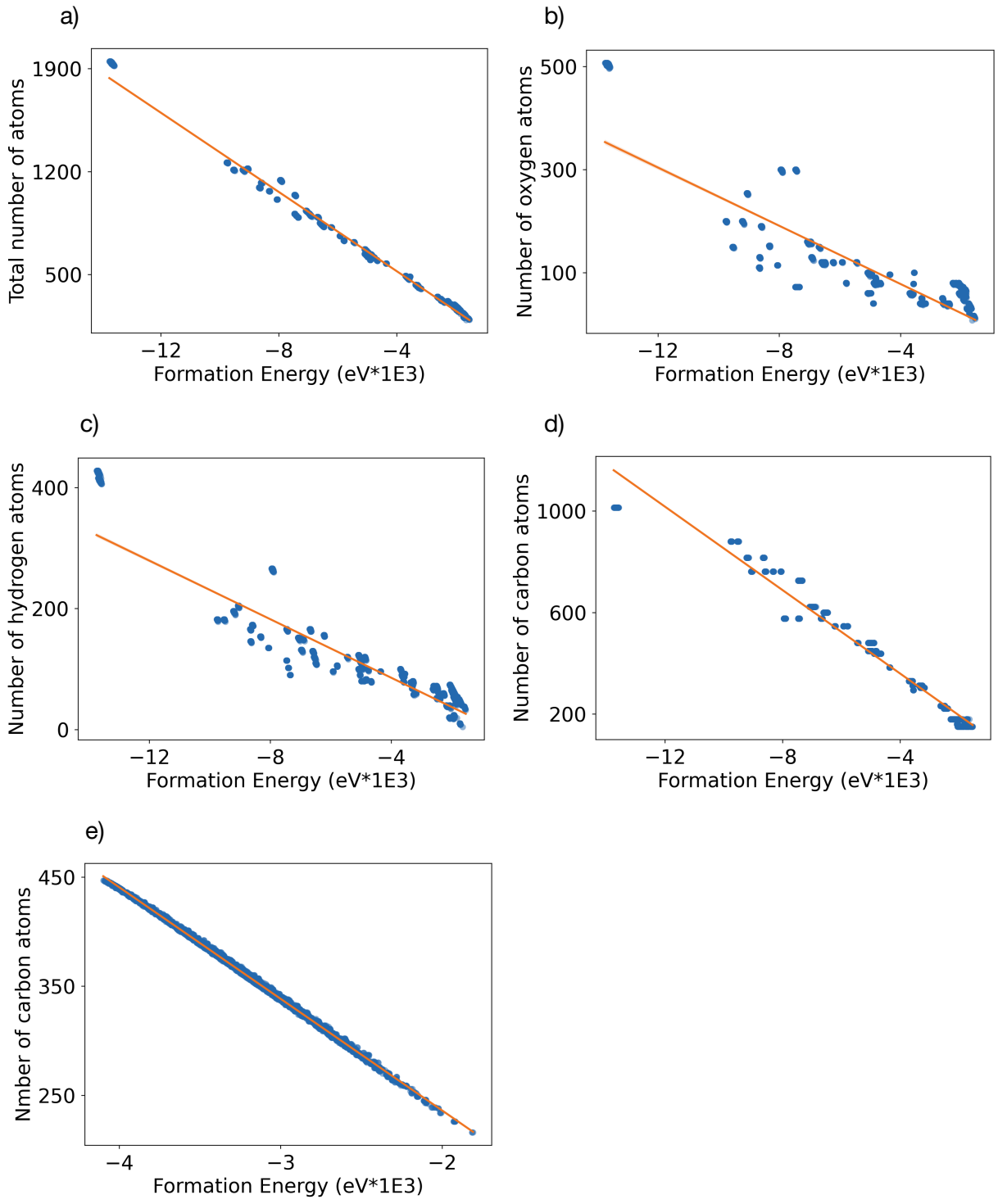


Figure S6. Distribution of formation energy for GO (a-d) and DG (e) and linear fits as a function of composition in the respective datasets.

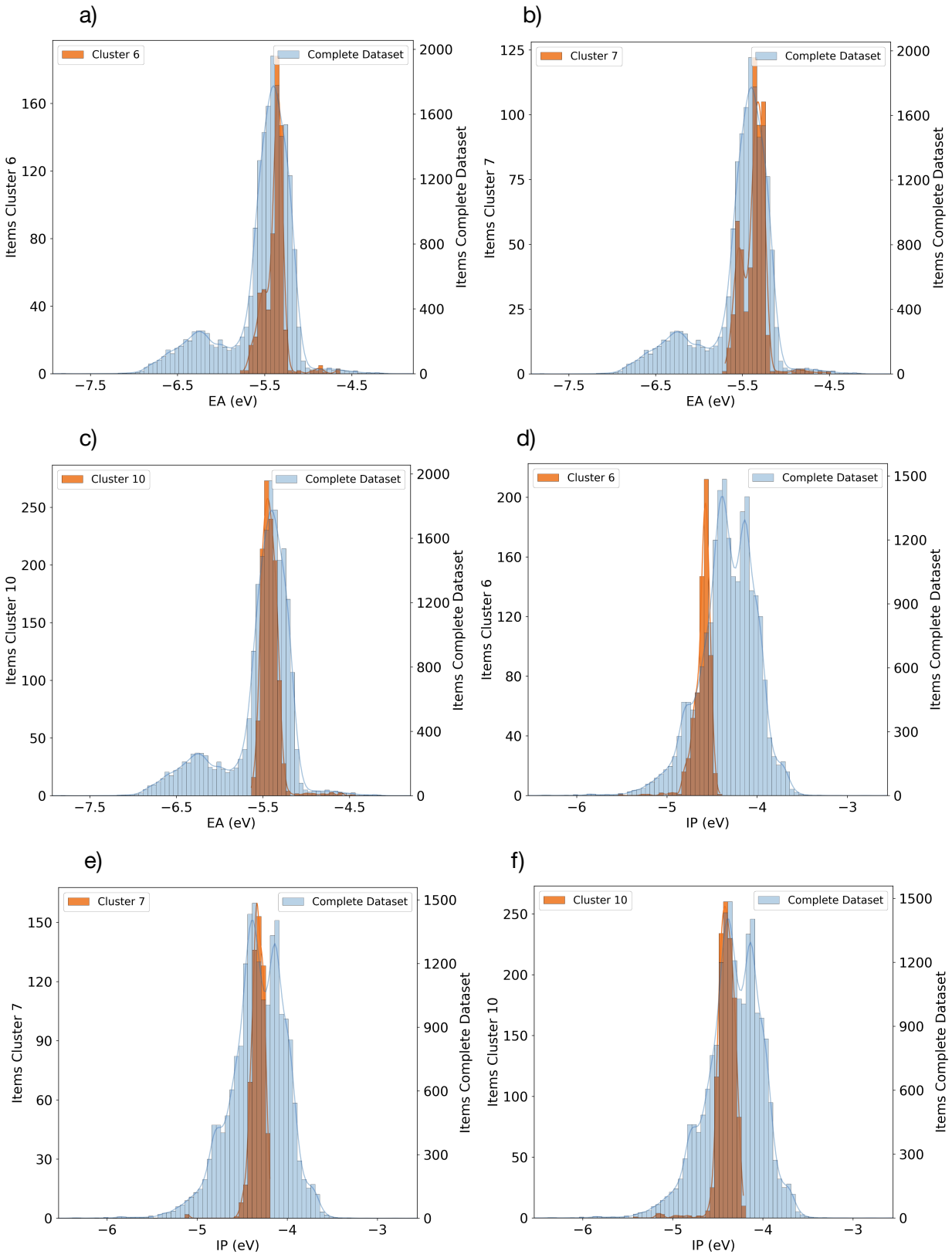


Figure S7. Distributions of EA (a-c) and IP (d-f) targets of the three selected UMAP clusters, namely cluster 6, 7 and 10, with respect to the complete dataset.

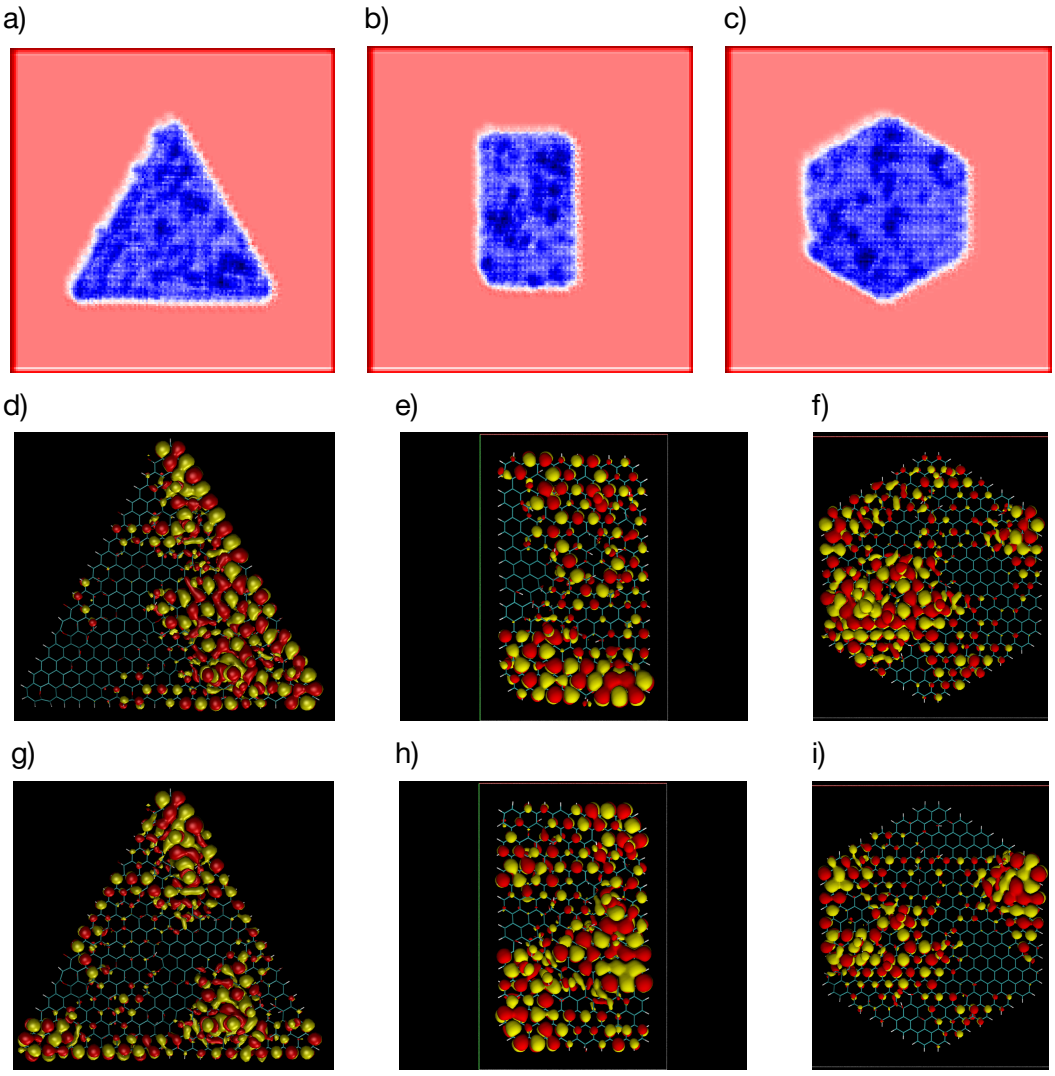
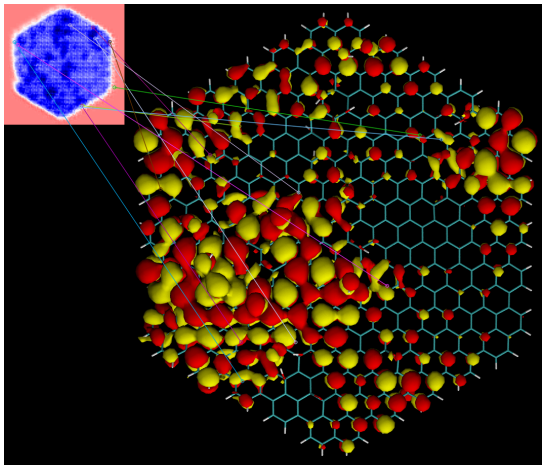
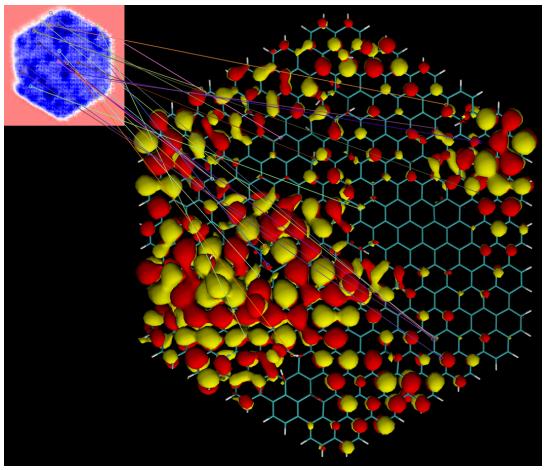


Figure S8. Feature maps for the high-level layers (layer 5, convolution 64, filter 62) (a-c); topology of HOMO (d-f) and LUMO (g-i) orbitals for selected GO samples with triangular (left), square (middle) and hexagonal (right) geometry, rendered with VMD[1].

a)



b)



c)

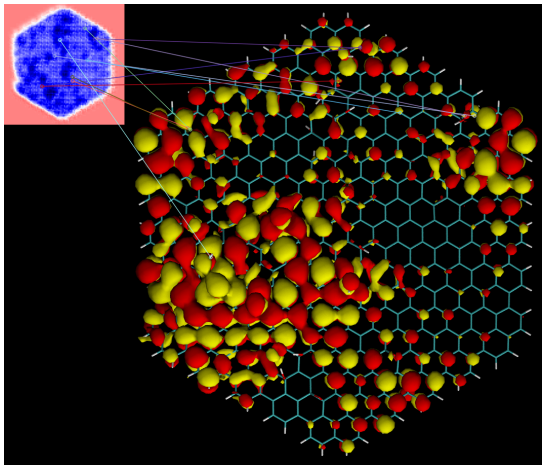


Figure S9. Correlation between IP features in image representations of a selected GO sample and HOMO topologies, rendered with VMD[1], obtained with the algorithms a) BRISK[5], b) FAST[6] and c) ORB[7].

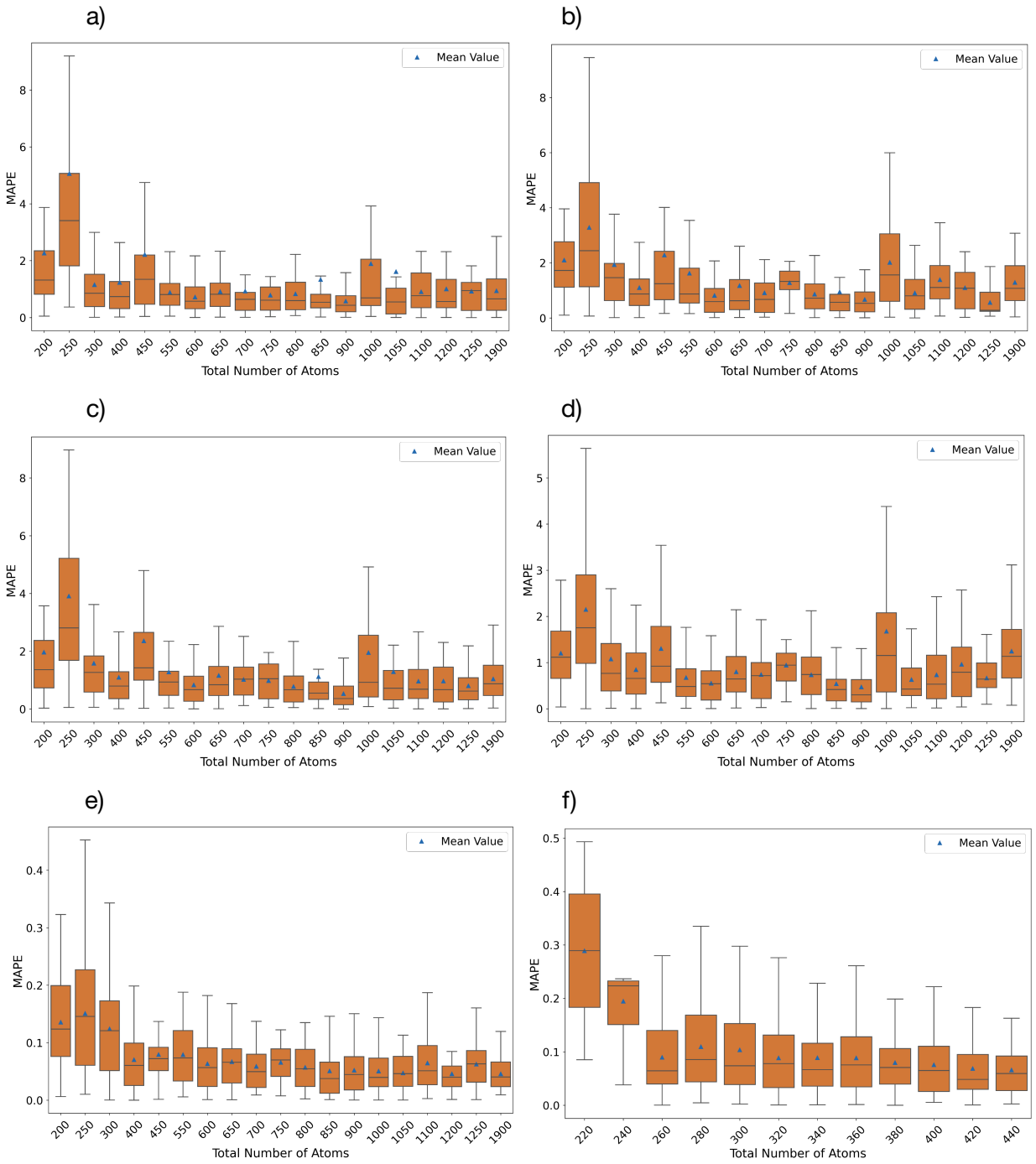


Figure S10. Boxplots illustrating the distribution, as a function of the total number of atoms in the analyzed samples, of the MAPE error of: a) electron affinity, b) ionization potential, c) electronegativity, d) Fermi energy and e) formation energy per atom for the GO reference dataset and f) formation energy per atom for DG reference dataset. The data are grouped into bins of 50 atoms each for GO and 20 atoms each for DG. Each boxplot displays the variability of MAPE within each bin, and the triangular marker represents the mean MAPE value for each bin.

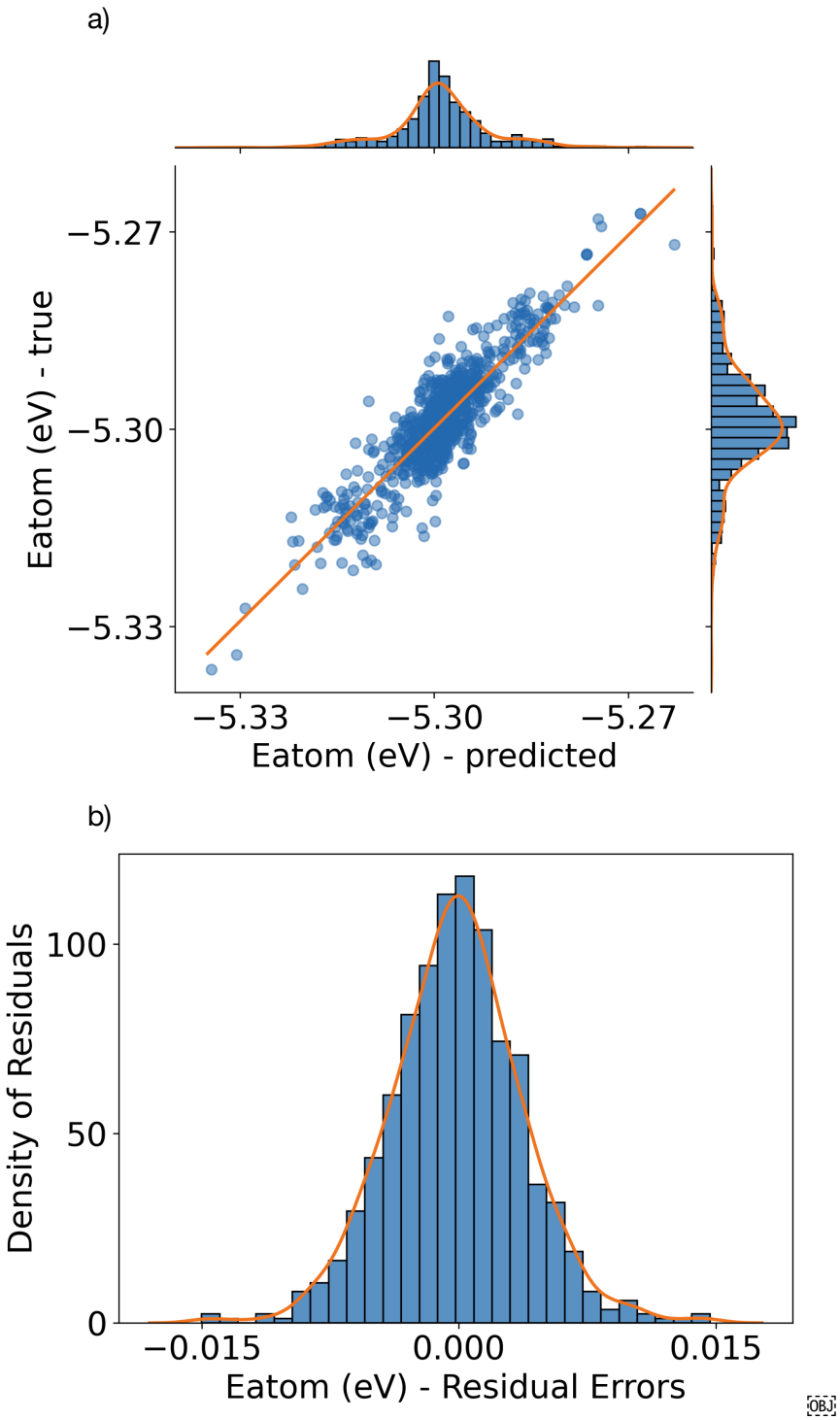


Figure S11. GrapheNet prediction for the defected phosphorene dataset for the total energy per atom target in terms of a) fit curve (R^2 = 0.794) and b) residual errors. Numerically, the mean, max and standard deviation values for the MAE (MAPE) errors are respectively 0.003 eV (0.056%), 0.015 eV (0.284%), 0.002 eV (0.047%) with a training time of 4.289 minutes.

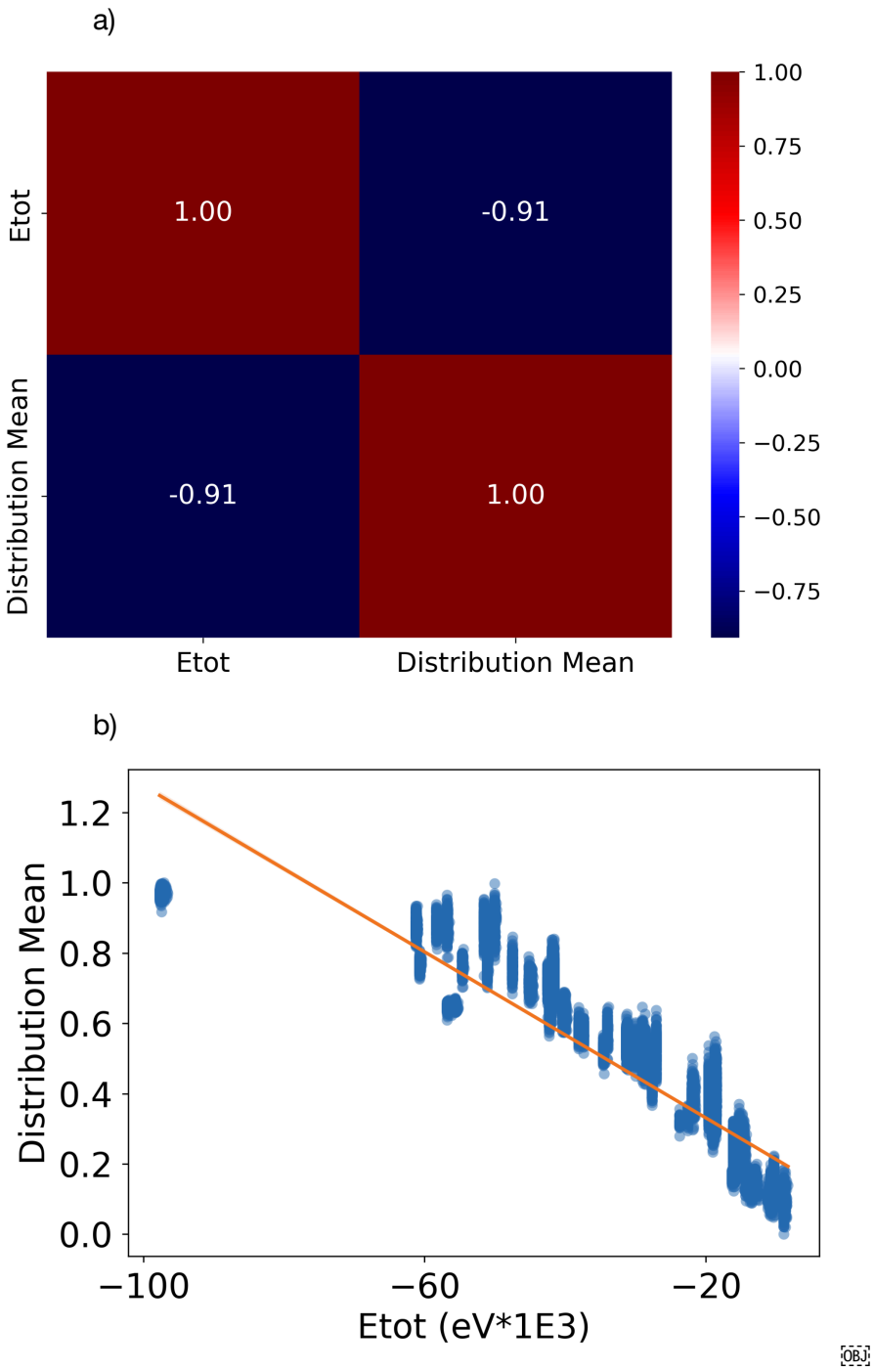


Figure S12. Correlation between total energy and the local distribution of oxygen for GO samples, represented as a) heatmap correlation, and b) regplot correlation. The local distribution is the normalized value that expresses the average of the distance between an oxygen atom and all other oxygen atoms.

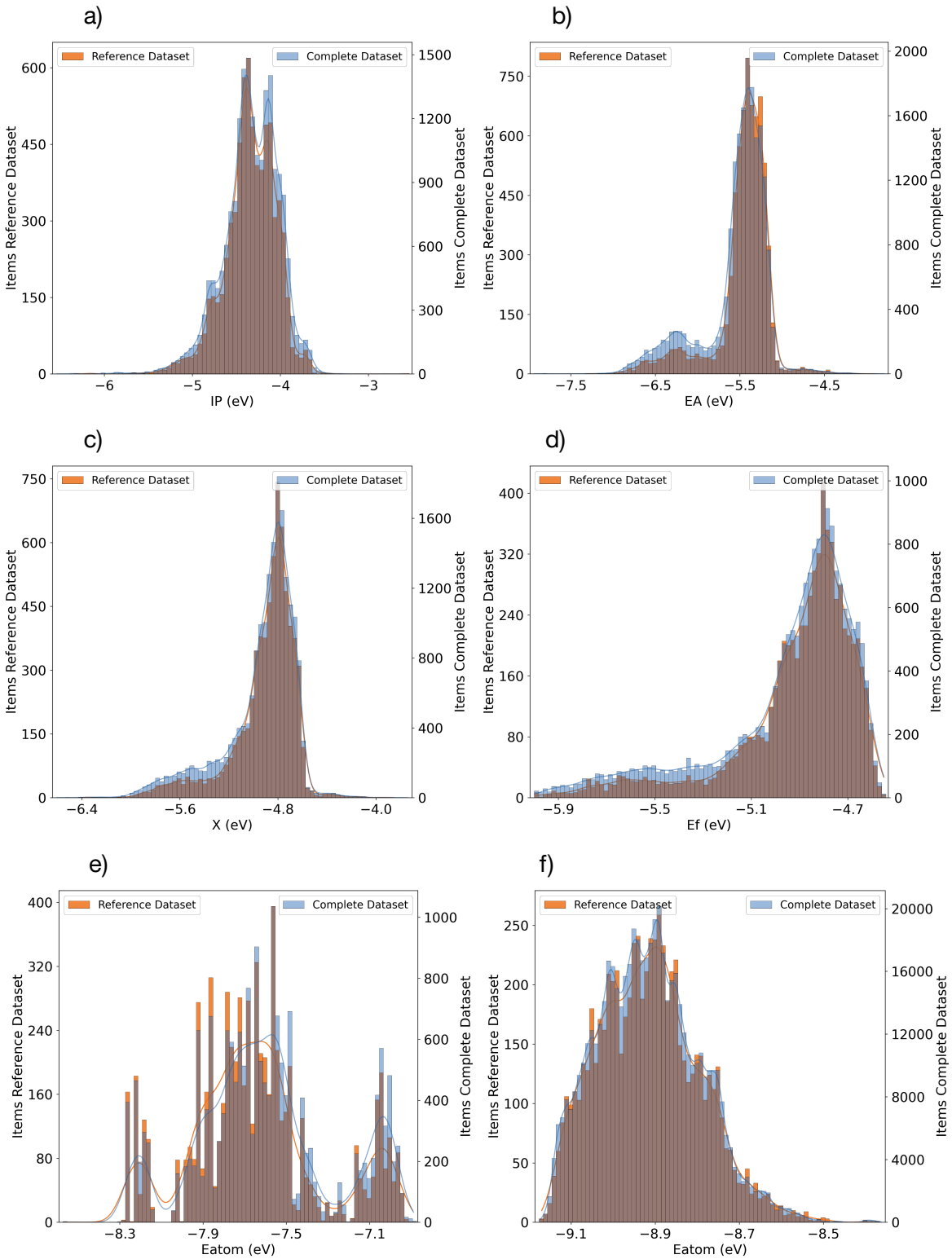


Figure S13. Distributions of a) ionization potential, b) electron affinity, c) electronegativity, d) Fermi energy and e) formation energy per atom of the reference GO dataset and f) formation energy per atom of the reference DG dataset, with respect their correspondent complete initial datasets.

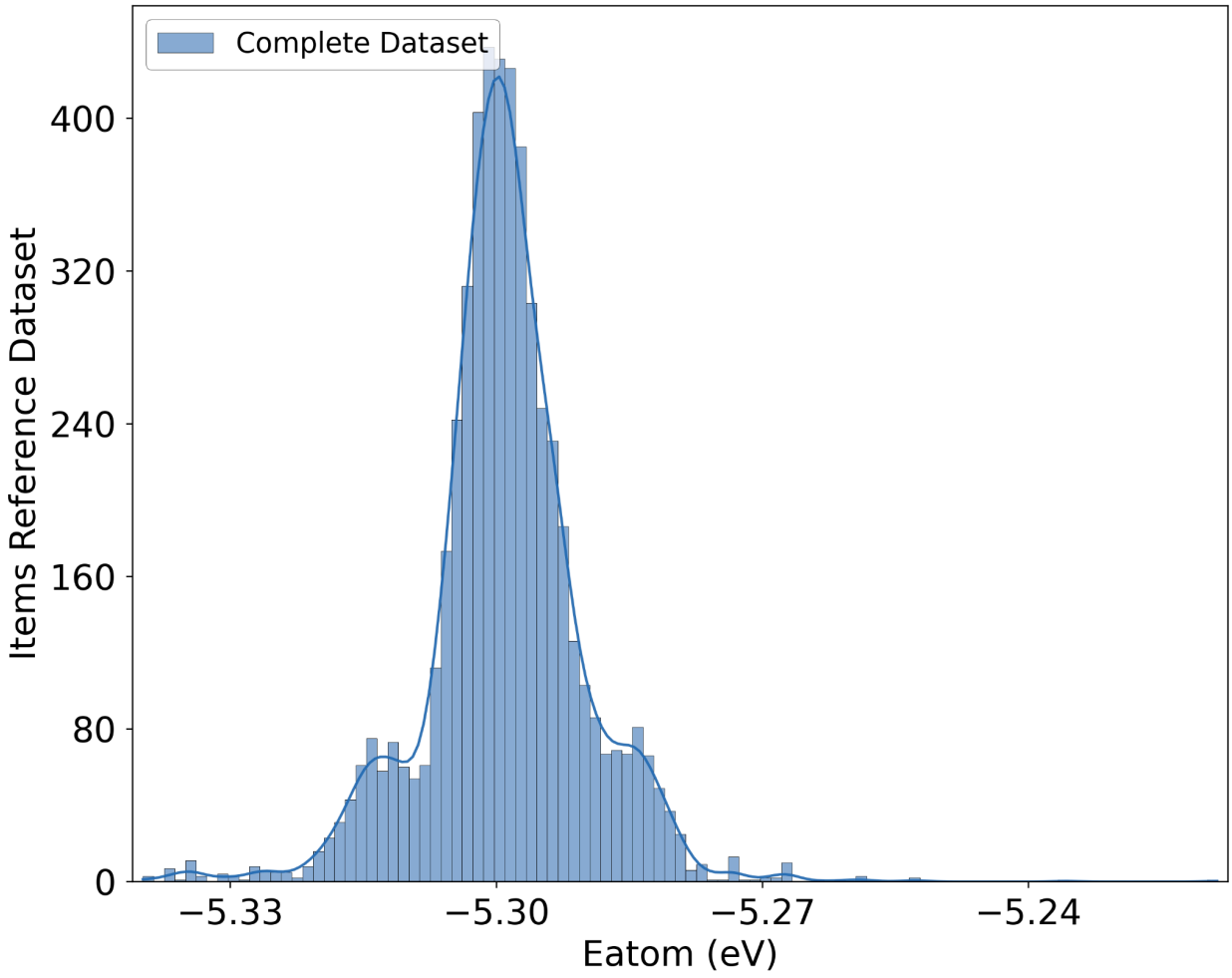


Figure S14. Distribution of the total energy per atom target of the complete dataset of defected phosphorene.

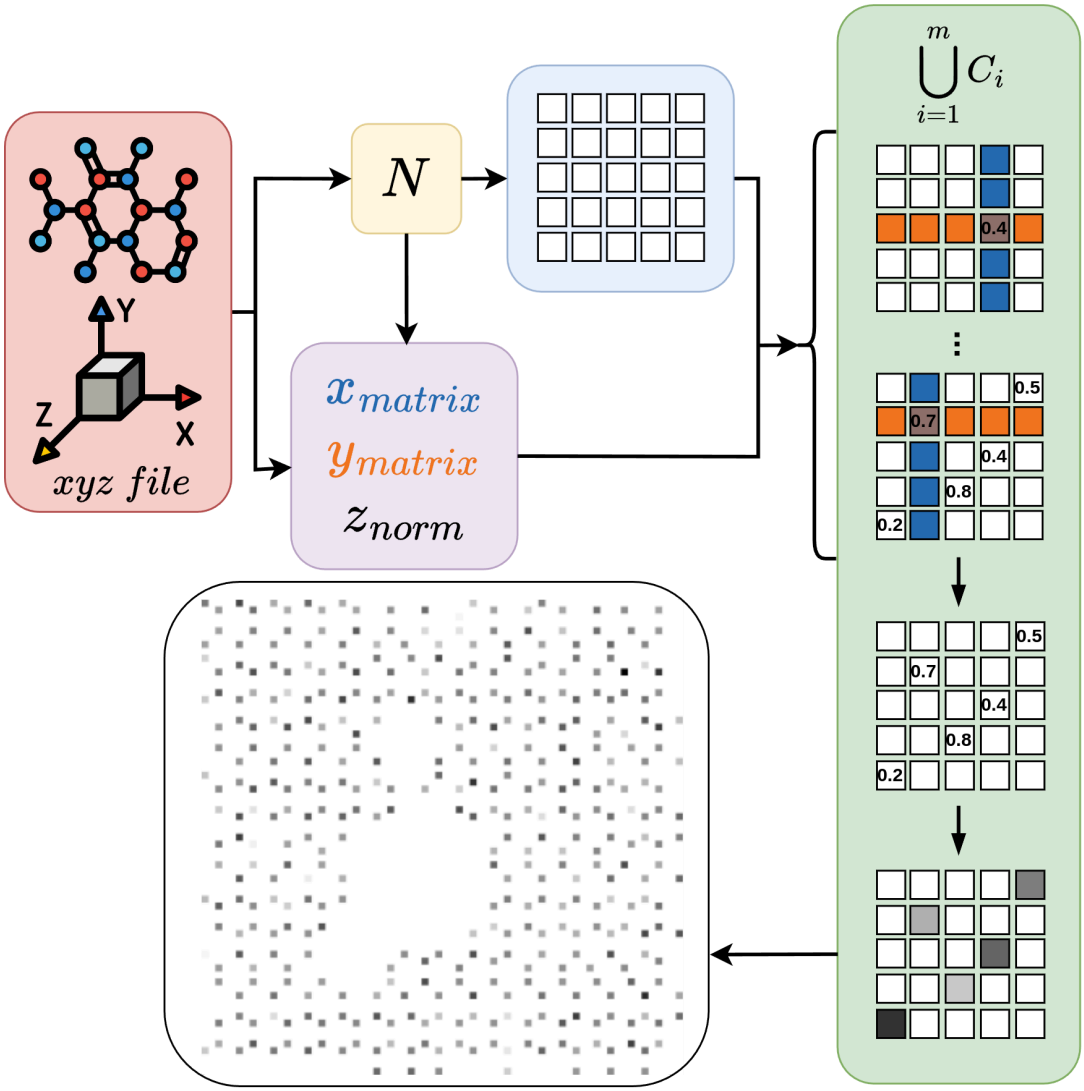


Figure S15. DG image generation workflow, plotted with drawio[4]. Starting from an xyz file, the factor *N* is computed using the maximum and minimum of the *x* and *y* coordinates. Then an empty *NxN* tensor is created. For each atom coordinate in the xyz file, the x_{matrix} and y_{matrix} coordinates of the tensor are computed, using the N factor. The indexed cell of the tensor is then filled with the maximum normalized *z* value (z_{norm}). This process is iterated over all the atoms in the xyz file, in order to completely populate the tensor. The tensor is then multiplied cell-wise by 255 and casted to 8-bit unsigned integer in order to be compatible with an image-like representation. Lastly, the tensor is converted into image and cropped using standard image manipulation libraries. The cropping part is used to remove all the portions of the image that do not contain relevant information, that is, all the black pixels around the structure, in order to minimize the size of the images themselves.

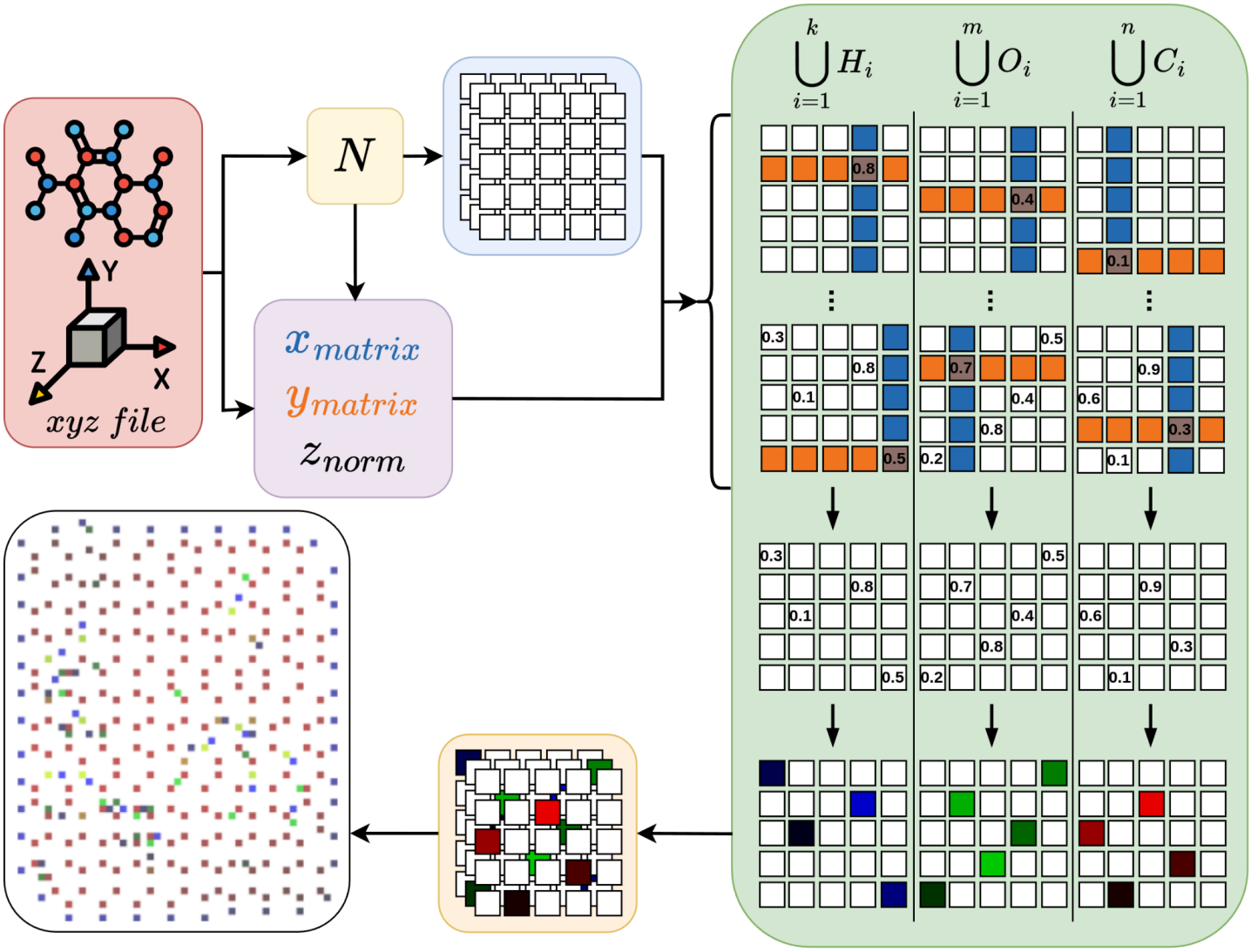


Figure S16. GO image generation workflow, plotted with drawio[4]. In the presence of multiple atom types, this contrast is depicted by utilizing distinct tensors that correspond to distinct color channels in the image, employing RGB instead of grayscale.

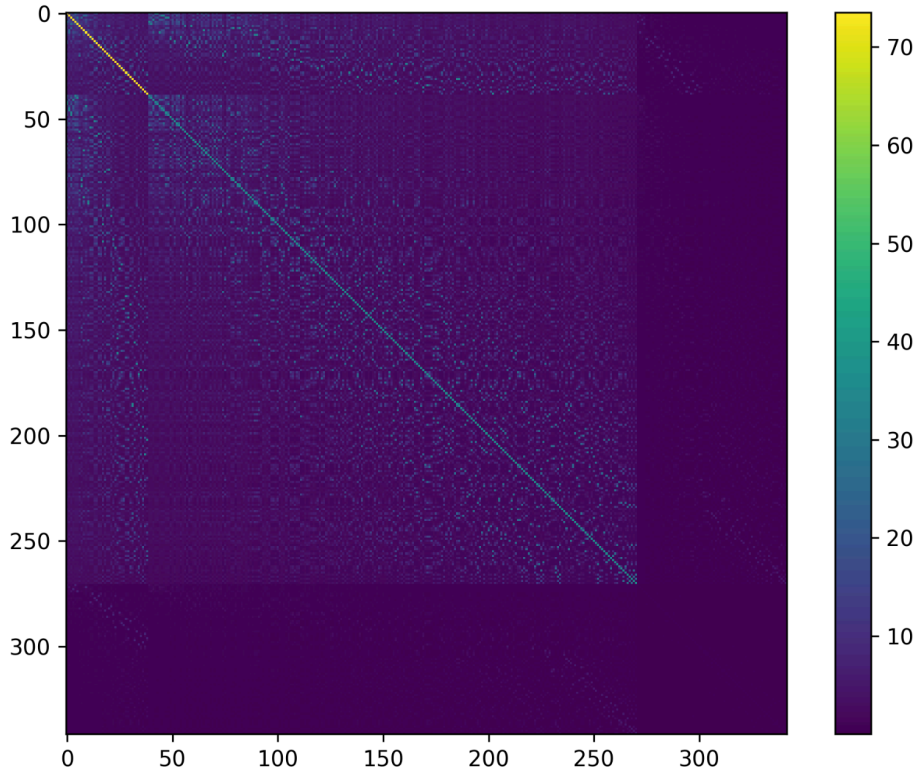


Figure S17. Representation of the Coulomb matrix for a selected GO sample.

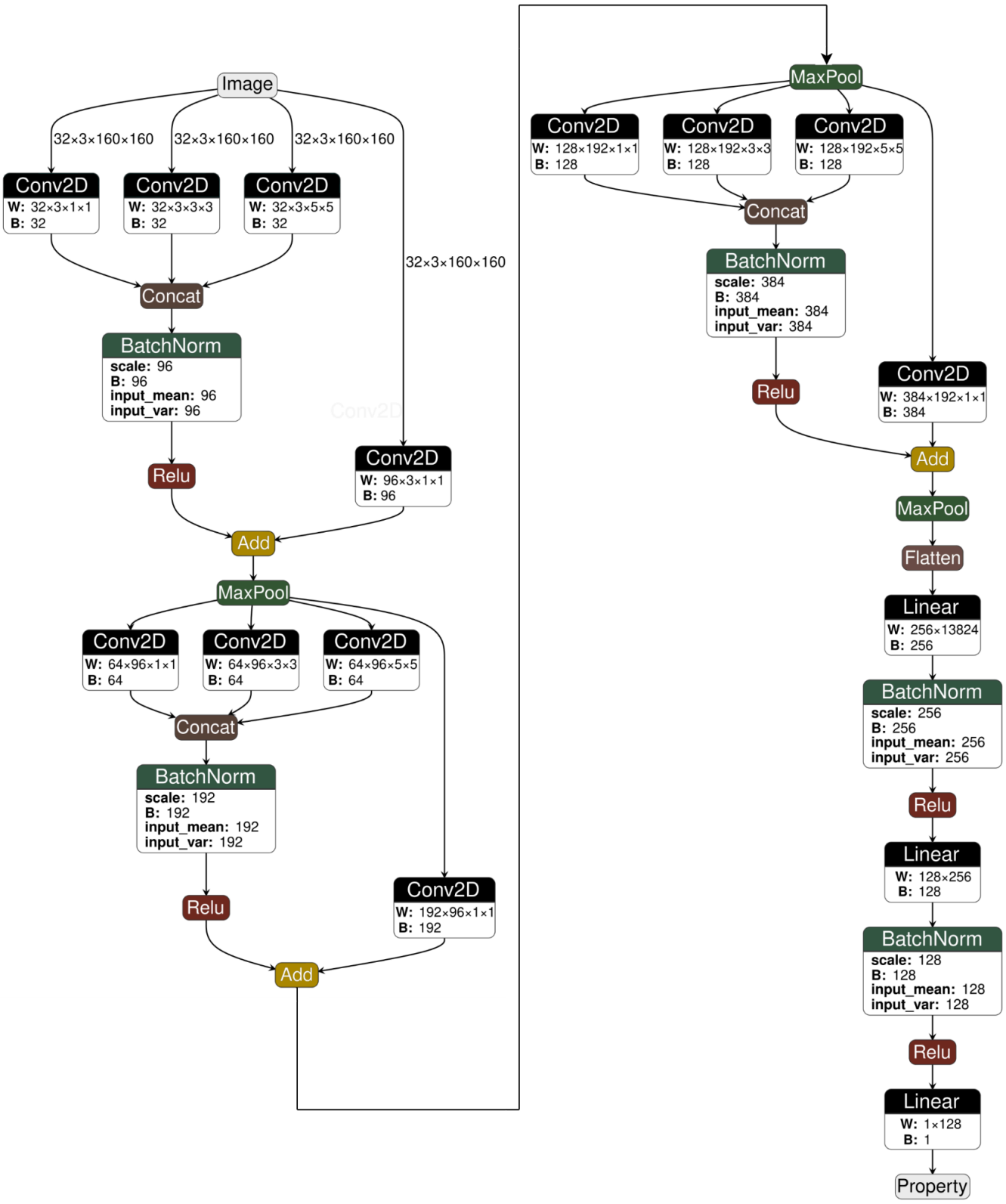


Figure S18. Inception-Resnet complete architecture with layers, dimensions and operations, plotted with Netron[8].

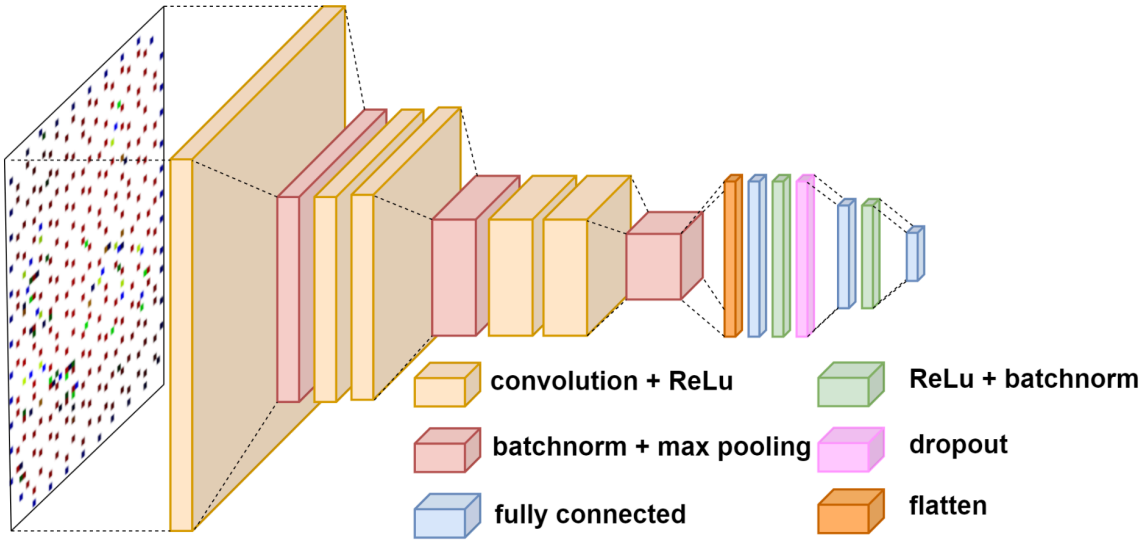


Figure S19. Scheme of the custom CNN network used for comparison with GrapheNet, plotted with drawio[4]. The custom CNN network consists of three blocks, each comprising consecutive convolutional and ReLU layers, with batch normalization and max pool layers between them. The resulting feature maps undergo flattening and pass through two fully connected layers with batch normalization and ReLU activation, followed by a dropout layer and a final fully connected layer for output predictions.

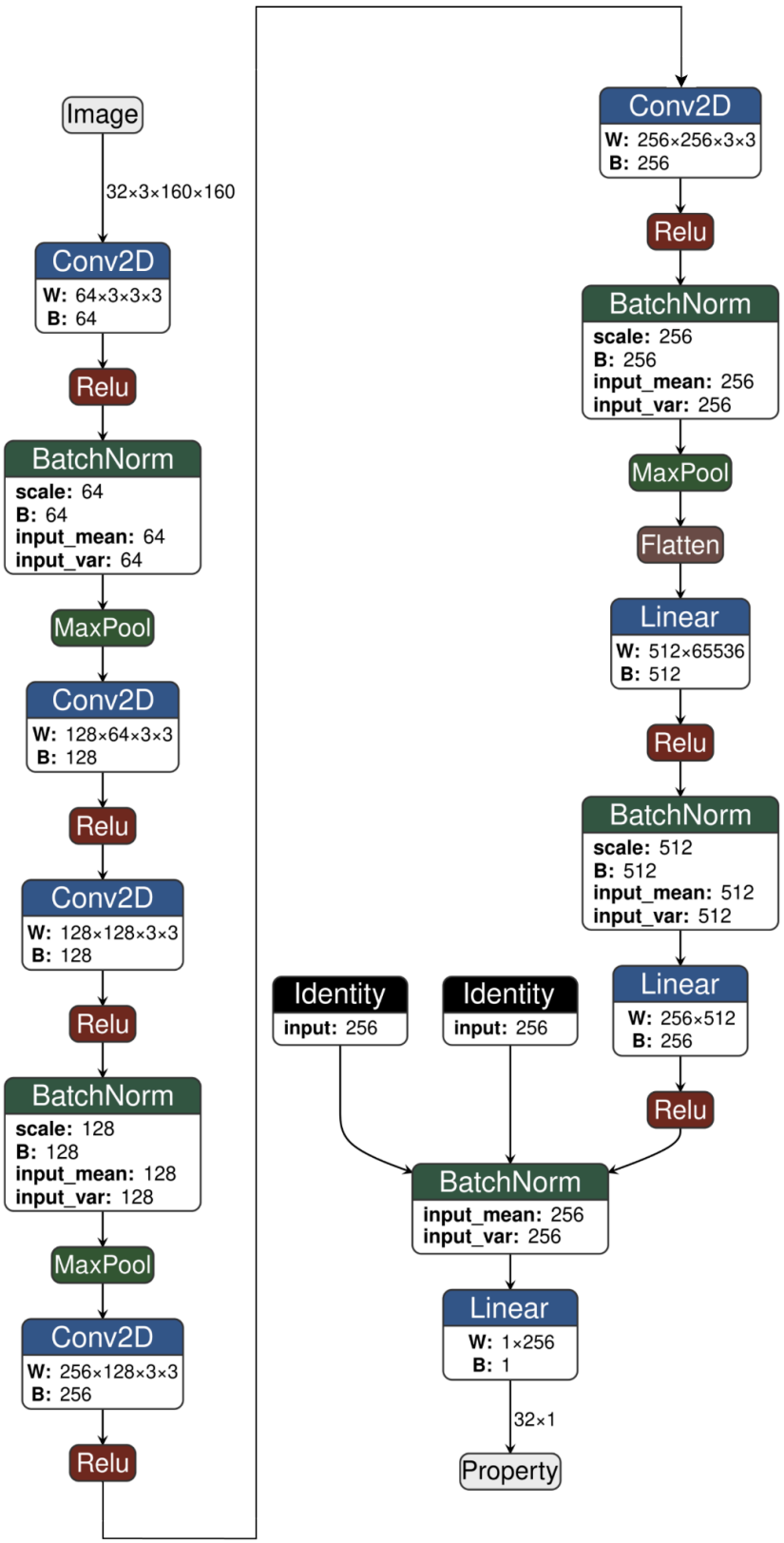


Figure S20. Custom CNN complete architecture with layers, dimensions and operations, plotted with Netron[8].

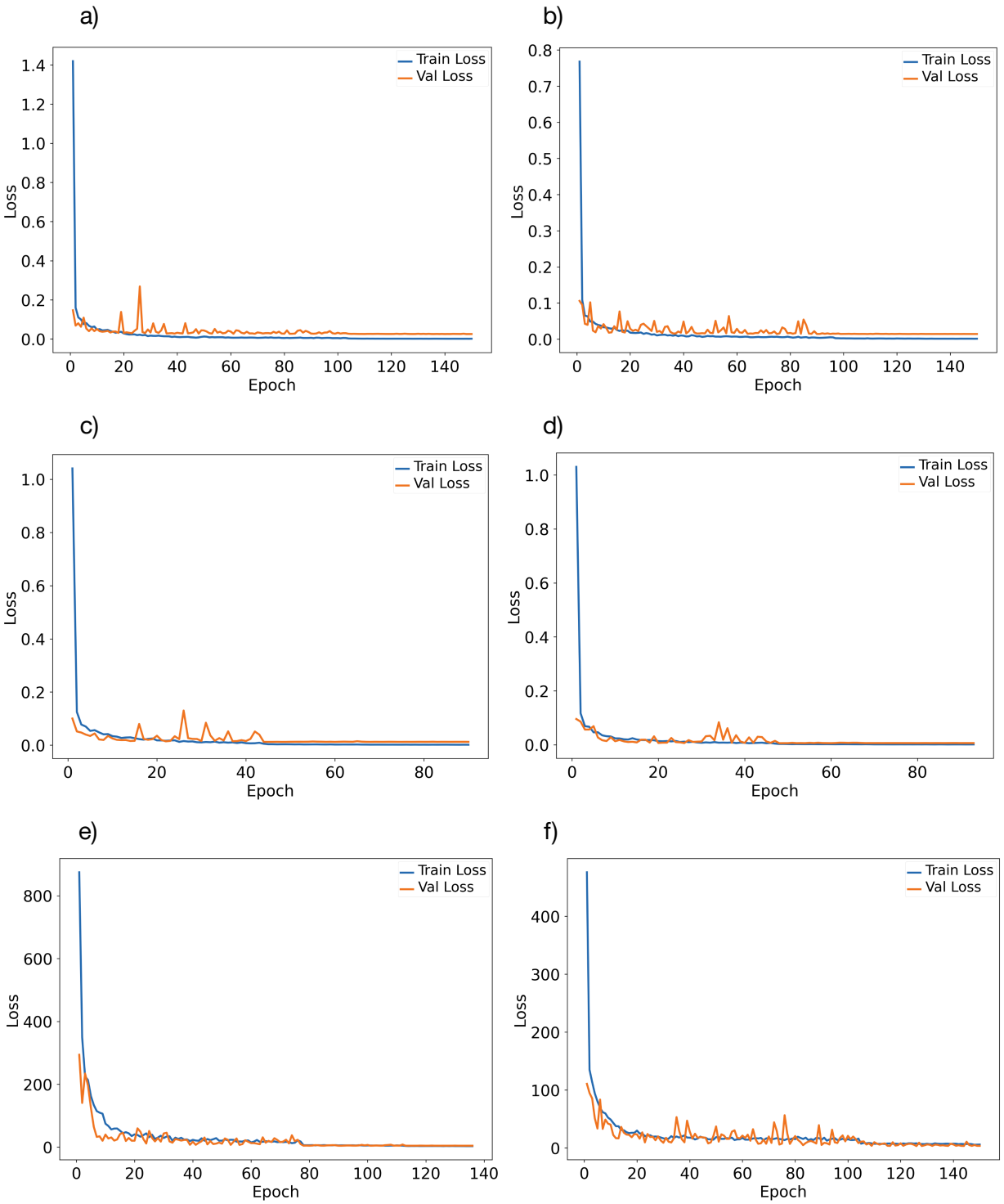


Figure S21. Learning curves, in terms of training loss (blue) and validation loss (orange) for: a) electron affinity, b) ionization potential, c) electronegativity, d) Fermi energy and e) formation energy per atom for the GO dataset and f) formation energy per atom for DG dataset.

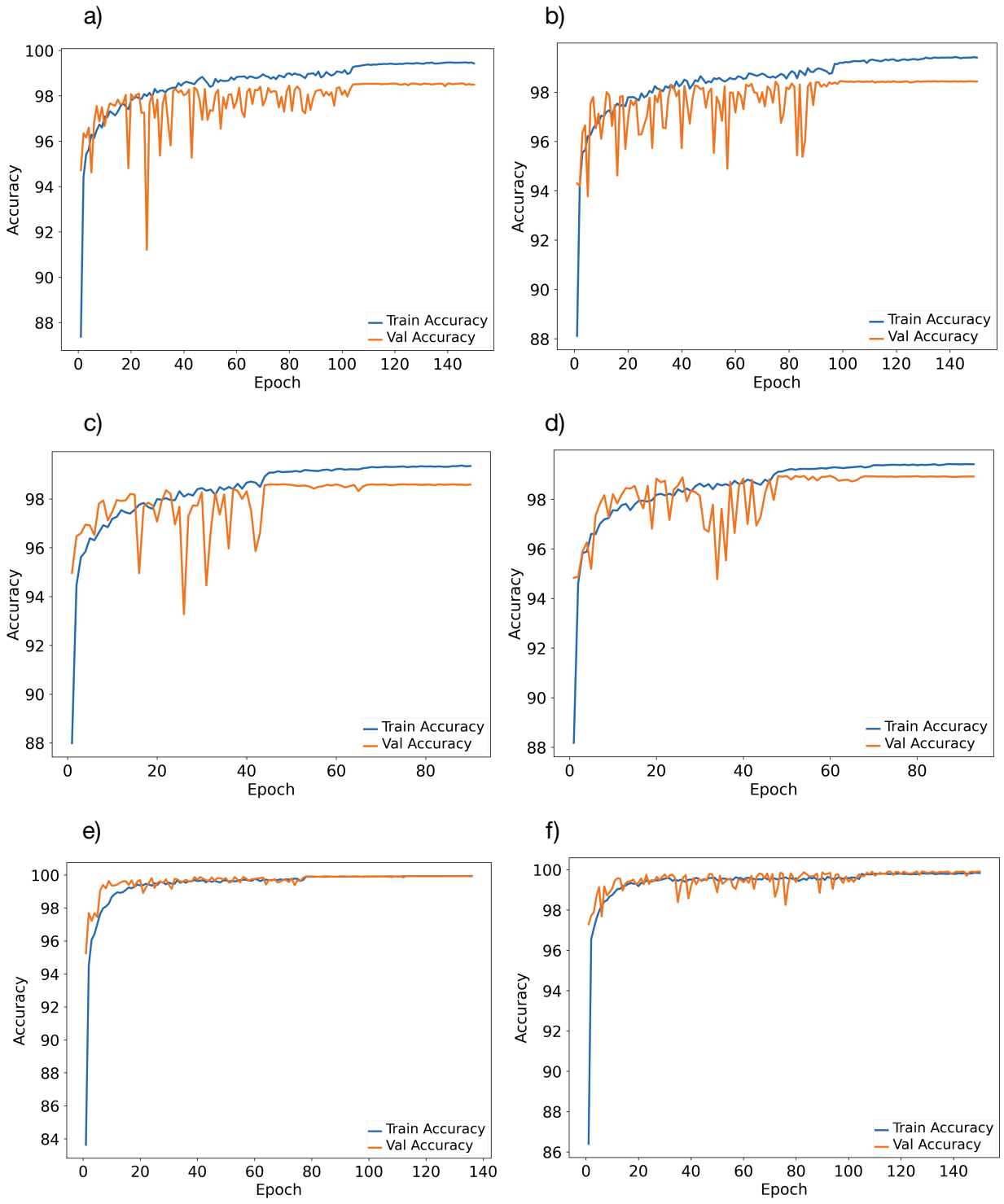


Figure S22. Learning curves, in terms of training accuracy (blue) and validation accuracy (orange): a) electron affinity, b) ionization potential, c) electronegativity, d) Fermi energy and e) formation energy per atom for the GO dataset and f) formation energy per atom for DG dataset.

Supplementary Tables

Table S1. Prediction results as maximum MAE (MAPE) error, mean MAE (MAPE) error and MAE (MAPE) standard deviation and training times (minutes) for the Custom CNN and Resnet18 models for the GO reference dataset. IP: ionization potential; EA: electron affinity; Ef: Fermi energy; χ: electronegativity, Eatom: formation energy per atom.

		Max	Mean	STD	Training time
CNN	IP	0.86 (17.64%)	0.07 (1.67%)	0.09 (1.99%)	25.27
	EA	1.50 (32.02%)	0.08 (1.52%)	0.15 (2.83%)	38.19
	Ef	0.54 (11.14%)	0.05 (1.01%)	0.06 (1.07%)	38.08
	χ	1.14 (26.36%)	0.08 (1.60%)	0.10 (1.97%)	19.64
	Eatom	0.03 (0.43%)	0.01 (0.08%)	0.01 (0.07%)	45.08
Resnet18	IP	0.91 (17.06%)	0.08 (1.70%)	0.10 (2.14%)	9.09
	EA	1.60 (34.07%)	0.09 (1.73%)	0.14 (2.83%)	11.78
	Ef	0.51 (8.55%)	0.06 (1.10%)	0.06 (1.14%)	10.49
	χ	1.27 (29.41%)	0.07 (1.49%)	0.10 (2.01%)	9.42
	Eatom	0.04 (0.57%)	0.01 (0.09%)	0.01 (0.08%)	17.63

Table S2. Complete GrapheNet MAE (MAPE) error metrics and training times (minutes) for the GO and DG reference datasets with images encoding.

		Max	Mean	STD	Training time
GO	IP	0.91 (18.05%)	0.06 (1.43%)	0.09 (1.90%)	9.90
	EA	1.64 (34.23%)	0.08 (1.48%)	0.15 (2.88%)	9.74
	Ef	0.41 (6.85%)	0.05 (0.97%)	0.05 (0.98%)	6.16
	χ	1.01 (23.34%)	0.07 (1.41%)	0.09 (1.91%)	6.00
	Eatom	0.03 (0.45%)	0.01 (0.08%)	0.01 (0.07%)	9.19
DG	Eatom	0.04 (0.49%)	0.01 (0.09%)	0.01 (0.07%)	6.22

Table S3. Prediction MAE (MAPE) errors and training times (minutes) for the GO reference dataset trained with grayscale images encoding on the GrapheNet CNN.

	Max	Mean	STD	Training time
IP	0.85 (16.92%)	0.07 (1.65%)	0.10 (2.06%)	7.70
EA	1.61 (34.38%)	0.10 (1.73%)	0.15 (2.90%)	10.96
Ef	0.63 (10.77%)	0.06 (1.21%)	0.07 (1.24%)	10.98
χ	1.14 (26.53%)	0.08 (1.56%)	0.10 (1.99%)	11.10
Eatom	0.04 (0.52%)	0.01 (0.08%)	0.01 (0.08%)	10.72

Table S4. Comparison of the MAE (MAPE) errors and training times (minutes) between the proposed image representation approach and the Coulomb matrix representation on the sub-dataset with samples constituted by 310 to 350 atoms. The dataset generation time is the time required to generate the dataset starting from the xyz files. Total computing time is evaluated as the sum between the dataset generation time and the maximum training time among the targets. The "Estimated Total Size" represents the total GPU memory usage (MB) of the network for a single forward and backward pass, including the memory required for input data, intermediate activations, gradient computations, and the storage of model parameters, computed with torchinfo[9].

	GrapheNet (Images)		**GrapheNet (Coulomb Matrices)**	
	MAE (MAPE)	Training time	MAE (MAPE)	Training time
IP	0.07 (1.67%)	1.78	0.07 (1.56%)	10.68
EA	0.07 (1.29%)	2.68	0.06 (1.14%)	6.72
Ef	0.05 (0.98%)	2.74	0.04 (0.83%)	7.66
χ	0.07 (1.37%)	2.15	0.05 (1.08%)	7.24
Eatom	0.01 (0.10%)	3.75	0.01 (0.09%)	14.14
Estimated Total Size (MB)	466.56		11210.11	
Dataset generation time	0.23 (2278 samples)		0.58 (2278 samples)	
Total time	**3.98**		**14.72**	

Table S5. Comparison of MAPE (MAE) errors and training times (minutes) between the proposed image representation approach and the Coulomb matrix representation on the sub-dataset with samples constituted by 610 to 660 atoms.

	GrapheNet (Images)		**GrapheNet (Coulomb Matrices)**	
	MAE (MAPE)	Training time	MAE (MAPE)	Training time
IP	0.04 (0.87%)	3.75	0.04 (0.83%)	40.55
EA	0.05 (1.01%)	3.36	0.05 (1.03%)	20.04
Ef	0.03 (0.63%)	3.93	0.03 (0.57%)	33.24
χ	0.05 (0.92%)	2.75	0.04 (0.75%)	25.72
Eatom	0.005 (0.06%)	4.47	0.005 (0.06%)	41.62
Estimated Total Size (MB)	1438.83		39841.72	
Dataset generation time	0.30 (1643 samples)		2.45 (1643 samples)	
Total time	**4.77**		**44.07**	

Table S6. Comparison of MAPE (MAE) errors and training times (minutes) between the proposed image representation approach and the Coulomb matrix representation on the sub-dataset with samples constituted by 310 to 660 atoms.

	GrapheNet (Images)		**GrapheNet (Coulomb Matrices)**	
	MAE (MAPE)	Training time	MAE (MAPE)	Training time
IP	0.06 (1.33%)	4.74	0.05 (1.22%)	100.08
EA	0.07 (1.27%)	5.61	0.06 (1.17%)	134.12
Ef	0.04 (0.81%)	5.91	0.04 (0.77%)	163.98
χ	0.05 (1.15%)	8.67	0.05 (1.14%)	65.92
Eatom	0.005 (0.07%)	9.02	0.006 (0.08%)	163.56
Estimated Total Size (MB)	1438.83		39841.72	
Dataset generation time	1.08 (7485 samples)		3.70 (7485 samples)	
Total time	**10.10**		**167.68**	

Table S7. Comparison of MAE (MAPE) errors and training times (minutes) between Kernel Ridge Regression, XGBoost and DNN on the sub-dataset with samples constituted by 310 to 660 atoms and using a representation based on the eigenvalues of the Coulomb matrix.

	KRR (Coulomb eigenvalues)		**XGBoost (Coulomb eigenvalues)**		**DNN (Coulomb eigenvalues)**	
	MAE (MAPE)	Training time	MAE (MAPE)	Training time	MAE (MAPE)	Training time
IP	0.06 (1.46%)	0.081	0.05 (1.18%)	0.084	0.06 (1.39%)	4.035
EA	0.07 (1.42%)	0.081	0.07 (1.26%)	0.083	0.07 (1.36%)	3.622
Ef	0.05 (1.07%)	0.156	0.04 (0.73%)	0.093	0.05 (0.97%)	3.538
χ	0.06 (1.31%)	0.083	0.05 (1.03%)	0.085	0.06 (1.25%)	3.199
Eatom	0.054 (0.71%)	0.080	0.009 (0.12%)	0.082	0.008 (0.10%)	3.656
Dataset generation time	29.250 (7485 samples)					
Total time	**29.406**		**29.343**		**33.285**	

Table S8. Comparison of MAE (MAPE) errors and training times (minutes) between Kernel Ridge Regression, XGBoost, DNN and GrapheNet on the sub-dataset with samples constituted by 1000 to 2000 atoms and using a representation based on the eigenvalues of the Coulomb matrix for KRR, XGBoost and DNN models, and on images for GrapheNet.

	KRR (Coulomb eigenvalues)		**XGBoost (Coulomb eigenvalues)**		**DNN (Coulomb eigenvalues)**		**GrapheNet (Images)**	
	MAE (MAPE)	Training time	MAE (MAPE)	Training time	MAE (MAPE)	Training time	MAE (MAPE)	Training time
IP	0.06 (1.22%)	0.036	0.05 (1.07%)	0.075	0.06 (1.19%)	3.731	0.06 (1.20%)	2.958
EA	0.06 (1.18%)	0.036	0.06 (1.11%)	0.074	0.06 (1.18%)	4.294	0.06 (1.15%)	7.000
Ef	0.05 (1.02%)	0.060	0.05 (0.88%)	0.112	0.05 (1.04%)	4.202	0.05 (0.97%)	6.786
χ	0.06 (1.13%)	0.036	0.05 (1.00%)	0.075	0.06 (1.12%)	4.132	0.06 (1.22%)	2.631
Eatom	0.026 (0.34%)	0.035	0.012 (0.16%)	0.066	0.018 (0.24%)	4.290	0.004 (0.06%)	6.947
Dataset generation time	77.500 (3183 samples)						1.900 (3183 samples)	
Total time	**77.560**		**77.612**		**81.794**		**8.900**	

Table S9. Comparison of MAPE (MAE) errors and training times (minutes) between the proposed image representation approach with GrapheNet and the graph matrix representation with M3GNet on the reference GO dataset.

	GrapheNet (Images)		**M3GNET (Graphs)**	
	MAE (MAPE)	Training time	MAE (MAPE)	Training time
IP	0.06 (1.43%)	9.90	0.06 (1.25%)	561.19
EA	0.08 (1.48%)	9.74	0.08 (1.45%)	424.91
Ef	0.05 (0.97%)	6.16	0.04 (0.70%)	783.80
χ	0.07 (1.41%)	6.00	0.05 (1.09%)	804.45
Eatom	0.01 (0.08%)	9.19	0.01 (0.07%)	805.93
Dataset generation time	1.10 (7000 samples)		8.10 (7000 samples)	
Total time	**11.00**		814.03	

Table S10. Minimum, maximum, mean and STD value (in eV) for each target in the reference datasets.

		Min	**Max**	**Mean**	**STD**
GO	IP	-6.34	-2.56	-4.34	0.29
	EA	-7.29	-3.93	-5.47	0.34
	Ef	-6.00	-4.55	-4.92	0.26
	χ	-6.40	-3.86	-4.91	0.27
	Eatom	-8.57	-6.89	-7.65	0.32
DG	Eatom	-9.17	-8.37	-8.91	0.12

Table S11. Minimum, maximum, mean and STD value (in eV) for each target in the complete initial datasets.

		Min	Max	Mean	STD
GO	IP	-6.54	-2.56	-4.33	0.32
	EA	-7.93	-3.80	-5.55	0.40
	Ef	-6.00	-4.55	-4.95	0.30
	χ	-6.60	-3.74	-4.94	0.30
	Eatom	-8.57	-6.89	-7.60	0.33
DG	Eatom	-9.18	-8.36	-8.92	0.12

Table S12. Comparison of MAE (MAPE) errors and training times (minutes) between different learning rates, namely 0.001, 0.01 (reference) and 0.1 on the reference datasets, with a fixed batch size of 32.

		LR=0.001		LR=0.01 (Reference)		LR=0.1	
		MAE (MAPE)	Training time	MAE (MAPE)	Training time	MAE (MAPE)	Training time
GO	IP	0.07 (1.56%)	12.277	0.06 (1.43%)	9.896	0.07 (1.64%)	5.203
	EA	0.10 (1.82%)	12.832	0.08 (1.48%)	9.735	0.09 (1.60%)	13.861
	Ef	0.06 (1.23%)	12.344	0.05 (0.97%)	6.159	0.05 (1.02%)	11.221
	χ	0.08 (1.56%)	7.206	0.07 (1.41%)	5.996	0.08 (1.60%)	5.495
	Eatom	0.082 (1.08%)	8.999	0.006 (0.08%)	9.187	0.006 (0.08%)	19.950
DG	Eatom	0.059 (0.67%)	2.299	0.008 (0.09%)	6.224	0.008 (0.09%)	4.000

Table S13. Comparison of MAE (MAPE) errors and training times (minutes) between different batch sizes, namely 16, 32 (reference) and 64 on the reference datasets, with a fixed learning rate of 0.01.

		BS=16		**BS=32 (Reference)**		**BS=64**	
		MAE (MAPE)	Training time	MAE (MAPE)	Training time	MAE (MAPE)	Training time
GO	IP	0.06 (1.45%)	9.627	0.06 (1.43%)	9.896	0.07 (1.53%)	6.790
	EA	0.08 (1.52%)	12.895	0.08 (1.48%)	9.735	0.09 (1.65%)	9.515
	Ef	0.05 (0.99%)	16.953	0.05 (0.97%)	6.159	0.05 (1.00%)	12.061
	χ	0.07 (1.37%)	12.774	0.07 (1.41%)	5.996	0.07 (1.35%)	9.873
	Eatom	0.006 (0.08%)	13.396	0.006 (0.08%)	9.187	0.007 (0.09%)	14.367
DG	Eatom	0.009 (0.10%)	7.363	0.008 (0.09%)	6.224	0.009 (0.10%)	4.085

Table S14. Comparison, in terms of bias and variance, between the results obtained by the proposed train/val/test split and the cross-validation with k-folds (6 and 12 folds) with images encoding on the GrapheNet CNN.

	Proposed		**6 folds**		**12 folds**	
	Bias	Variance	Bias	Variance	Bias	Variance
IP	2.70e-3	1.74e-2	3.22e-3	4.60e-3	3.35e-3	4.71e-3
EA	3.33e-3	9.77e-3	-8.54e-4	8.23e-3	-1.53e-3	8.13e-3
Ef	-2.67e-3	3.03e-2	4.71e-3	2.71e-2	5.19e-3	2.75e-2
χ	5.71e-3	1.35e-2	2.46e-3	8.31e-3	4.46e-4	8.57e-3
Eatom	-1.06e-4	5.69e-2	-1.34e-5	3.44e-2	1.01e-6	3.46e-2

References

[1] Humphrey, W., Dalke, A., and Schulten, K. ‘VMD -- Visual Molecular Dynamics’, Journal of Molecular Graphics, vol. 14, pp. 33–38, DOI: https://doi.org/10.1016/0263-7855(96)00018-5 (1996).
[2] Bradski, G. The OpenCV Library. Dr. Dobb’s J. Softw. Tools DOI: https://github.com/opencv/opencv (2000).
[3] Murray, A. et al. python-pillow/pillow: 10.4.0, DOI: https://doi.org/10.5281/zenodo.12606429 (2024)
[4] JGraph, ‘draw.io’. https://github.com/jgraph/drawio (2021).
[5] Leutenegger, S., Chli, M., & Siegwart, R.Y. "BRISK: Binary Robust invariant scalable keypoints", International Conference on Computer Vision, pp. 2548-2555, DOI: https://doi.org/10.1109/ICCV.2011.6126542 (2011).
[6] Rosten, E., Porter, R., & Drummond, T. “Faster and better: A machine learning approach to corner detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence, 32*(1), 105-119. DOI: https://doi.org/10.1109/TPAMI.2008.275 (2010).
[7] Rublee, E., Rabaud, V., Konolige, K. & Bradski, G. “ORB: an efficient alternative to SIFT or SURF”, Proceedings of the IEEE International Conference on Computer Vision, 2564-2571, DOI: https://doi.org/10.1109/ICCV.2011.6126544 (2011).
[8] Roeder, L. "Netron: Visualizer for neural network, deep learning and machine learning models.”, https://github.com/lutzroeder/netron.
[9] Yep, T. “torchinfo”, https://github.com/TylerYep/torchinfo/tree/main (2020).