Supplemental material: Unlocking 3D Nanoparticle Shapes from 2D High-Resolution Transmission Electron Microscopy images: A Deep Learning Breakthrough

Romain Moreau,[1] Hakim Amara,[1,2] Maxime Moreaud,[3,4] Jaysen Nelayah,[2] Adrien Moncomble,[2] Christian Ricolleau,[2] Damien Alloyeau,[2] Guillaume Wang,[2] and Riccardo Gatti[1]

[1] *Université Paris-Saclay, ONERA, CNRS, Laboratoire d'Étude des Microstructures (LEM), 92322, Châtillon, France*
[2] *Université Paris Cité, Laboratoire Matériaux et Phénomènes Quantiques (MPQ), CNRS-UMR7162, 75013 Paris, France*
[3] *IFP Energies nouvelles, Rond-point de l'échangeur de Solaize, BP 3, 69360 Solaize, France*
[4] *Manufacture Française des Pneumatiques Michelin, 23 Place des Carmes Déchaux, 63000 Clermont-Ferrand, France*

(Dated: July 18, 2025)

I. COMPARISON OF DIFFERENT ARCHITECTURES: WHY 6 CONVOLUTIONAL LAYERS IN THE PLAIN CNN?

In the early stages of this study, several architectures have been compared and a few significant trends have been identified (see Figure S1). As shown in the article, ResNet34 systematically outperforms "shallower" CNNs, no matter the size of the training dataset, and the gap is rather significant. The gain in accuracy lies between 2 to 5% depending on the size of the training dataset. It has also been proven that a CNN containing 6 blocks (Convolution+Pooling) is better than a CNN containing 5 blocks, which also performs better than a CNN comprising 4 blocks. When the CNN has 6 of these blocks, it allows for more down-sampling, with filters of size 2×2 just before flattening. This allows for a better abstraction of the input data, by making an interaction between pixels that would not interact with less blocks.

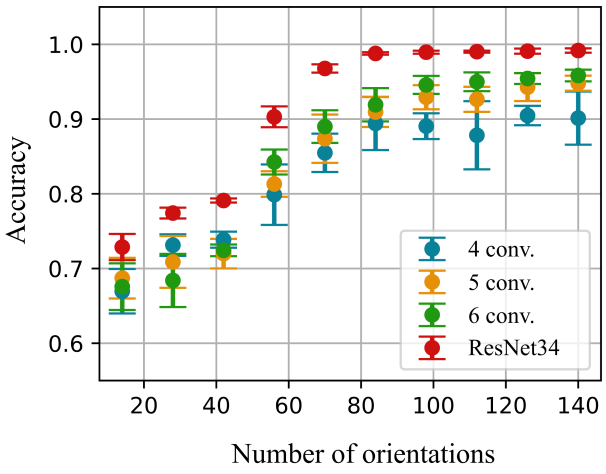

FIG. S1. The predictive accuracy of the classifier as a function of the size of the training dataset. This gives a comparison of performance depending on the architecture of the CNN model, with three "shallow" models containing 4, 5 and 6 convolutional layers, respectively, and one deeper model: ResNet34.

II. CONFUSION MATRICES CORRESPONDING TO THE OPTIMIZATION OF THE TRAINING DATASET

To have a deeper understanding of how accurate each model is, depending on the way the dataset is generated, it is possible to look at confusion matrices. Here, as seen in the article, one can see a competition between the sampling of defocus in the focal series and the number of orientations for each NP (see Figure S2).

Overall, cubes are very well disentangled from the four other shapes. The lowest prediction accuracy is consistently for cuboctahedral shapes, followed by truncated octahedra. The most common misclassifications occur for cuboctahedra, being predicted as icosahedra. This could be surprising given the fact that cuboctahedra, that can be seen as being halfway between a truncated cube and a truncated octahedron, should be mistaken for what we call truncated octahedra in this study, which is not the case. It could be explained by the fact that under some exotic orientations, which are allowed by our sampling of the orientations, cuboctahedra tend to look like icosahedra, locally. Overall, of course, such icosahedral shapes should exhibit some twinning that are not to be seen in any other shape in our study.

As the number of orientations increases in the training dataset, the overall accuracy also increases, but here it is possible to see that the gain in predictive accuracy increases simultaneously for cuboctahedra, truncated octahedra and octahedra, up to the point of reaching almost 100% for octahedra. Accuracy for cuboctahedra remains significantly lower than the other shapes. Despite reaching sometimes 93 to 94%, it can stall around 90%, as shown on Figure S2(b) on the two last matrices. Such fluctuations are explained by different weights and biases initialization.

III. ROTATION MATRICES

$R_{\eta \vec{u}}$ and $R_{\psi\theta\phi}$ are given explicitly in equation 1 and equation 2.

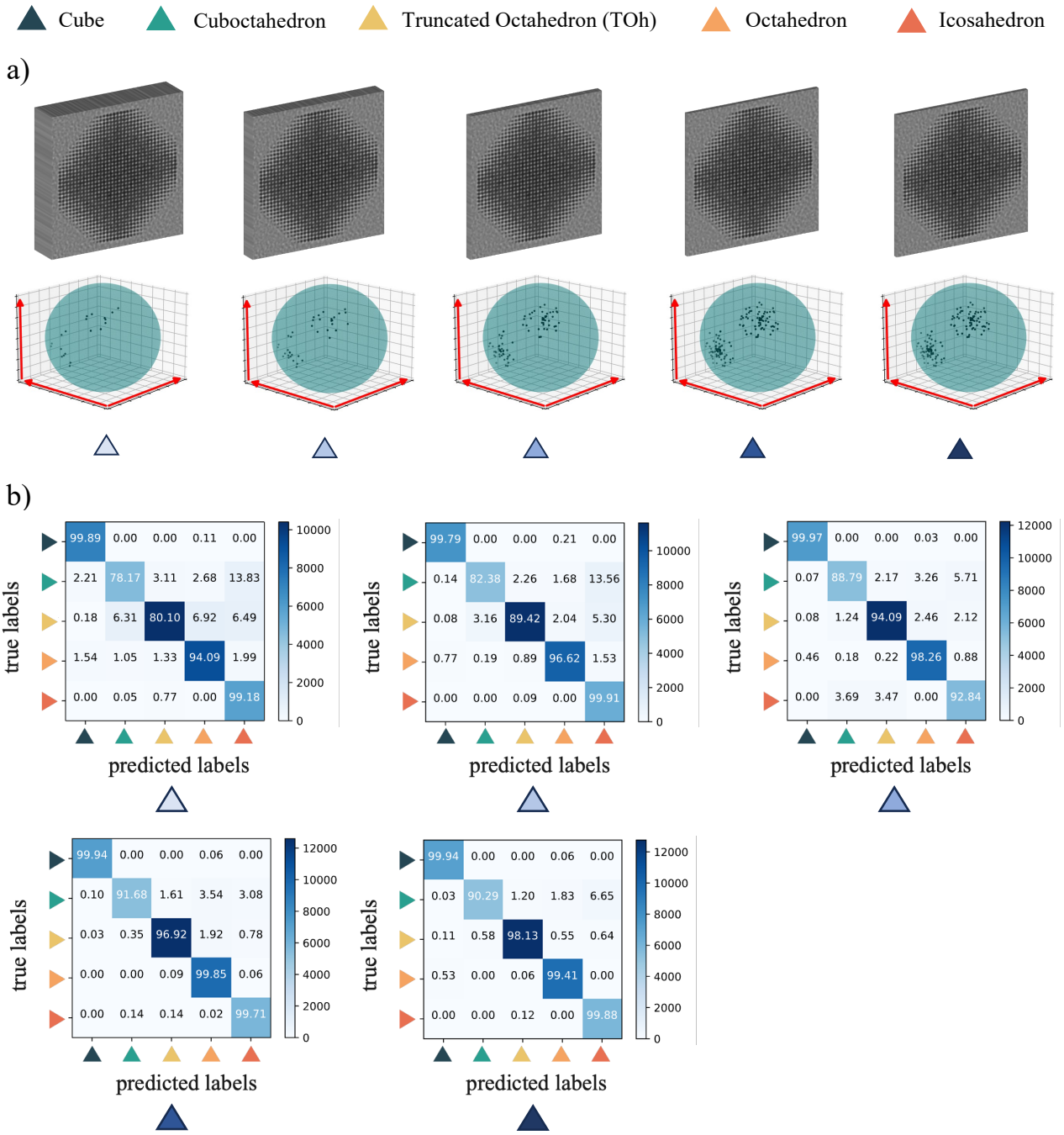

FIG. S2. a) Visualization of the sampling of focal series and the distribution of NP orientations on the 3D sphere. b) Examples of confusion matrices for each dataset, chosen after one of 10 independent trainings (with different weights and biases initialization).

IV. EXAMPLES OF NPS MORPHOLOGY ANALYSIS BASED ON THIS CNN APPROACH FROM HRTEM IMAGES

As shown in the article, this approach is very efficient for the prediction of the shape of NPs. Here, an large number of images are shown to emphasize the robustness of the model, by varying defocusing conditions, the orientation and the size of NPs. The CNN consistently predicts each shape accurately with a near 100% probability (see Figure S3).

$$R_{\eta\vec{u}} = \begin{pmatrix} \cos\eta + u_x^2\,(1-\cos\eta) & u_x u_y\,(1-\cos\eta) - u_z\sin\eta & u_x u_z\,(1-\cos\eta) + u_y\sin\eta \\ u_y u_x\,(1-\cos\eta) + u_z\sin\eta & \cos\eta + u_y^2\,(1-\cos\eta) & u_y u_z\,(1-\cos\eta) - u_x\sin\eta \\ u_z u_x\,(1-\cos\eta) - u_y\sin\eta & u_z u_y\,(1-\cos\eta) + u_x\sin\eta & \cos\eta + u_z^2\,(1-\cos\eta) \end{pmatrix} \tag{1}$$

$$\begin{aligned} R_{\psi\theta\phi} &= \begin{pmatrix} \cos(\psi) & \sin(\psi) & 0 \\ -\sin(\psi) & \cos(\psi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta) & \sin(\theta) \\ 0 & -\sin(\theta) & \cos(\theta) \end{pmatrix} \begin{pmatrix} \cos(\phi) & \sin(\phi) & 0 \\ -\sin(\phi) & \cos(\phi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ R_{\psi\theta\phi} &= \begin{pmatrix} \cos(\psi)\cos(\phi) - \cos(\theta)\sin(\phi)\sin(\psi) & \cos(\psi)\sin(\phi) + \cos(\theta)\cos(\phi)\sin(\psi) & \sin(\psi)\sin(\theta) \\ -\sin(\psi)\cos(\phi) - \cos(\theta)\sin(\phi)\cos(\psi) & -\sin(\psi)\sin(\phi) + \cos(\theta)\cos(\phi)\cos(\psi) & \cos(\psi)\sin(\theta) \\ \sin(\theta)\sin(\phi) & -\sin(\theta)\cos(\phi) & \cos(\theta) \end{pmatrix} \end{aligned} \tag{2}$$

V. EXAMPLES OF MISCLASSIFICATIONS IN THE TEST DATASET

A focus is made on the misclassifications, knowing that almost all the cases where the CNN fails are included here (see Figure S4). As discussed in the article, a significant amount of classification errors occur when the NPs are oriented in such a way that contrast is very low, the aspect of the NPs is blurred and the edges are not well defined. This effect is strengthened by poor defocusing conditions. The expected shape is often predicted as the second choice, with a rather high probability.

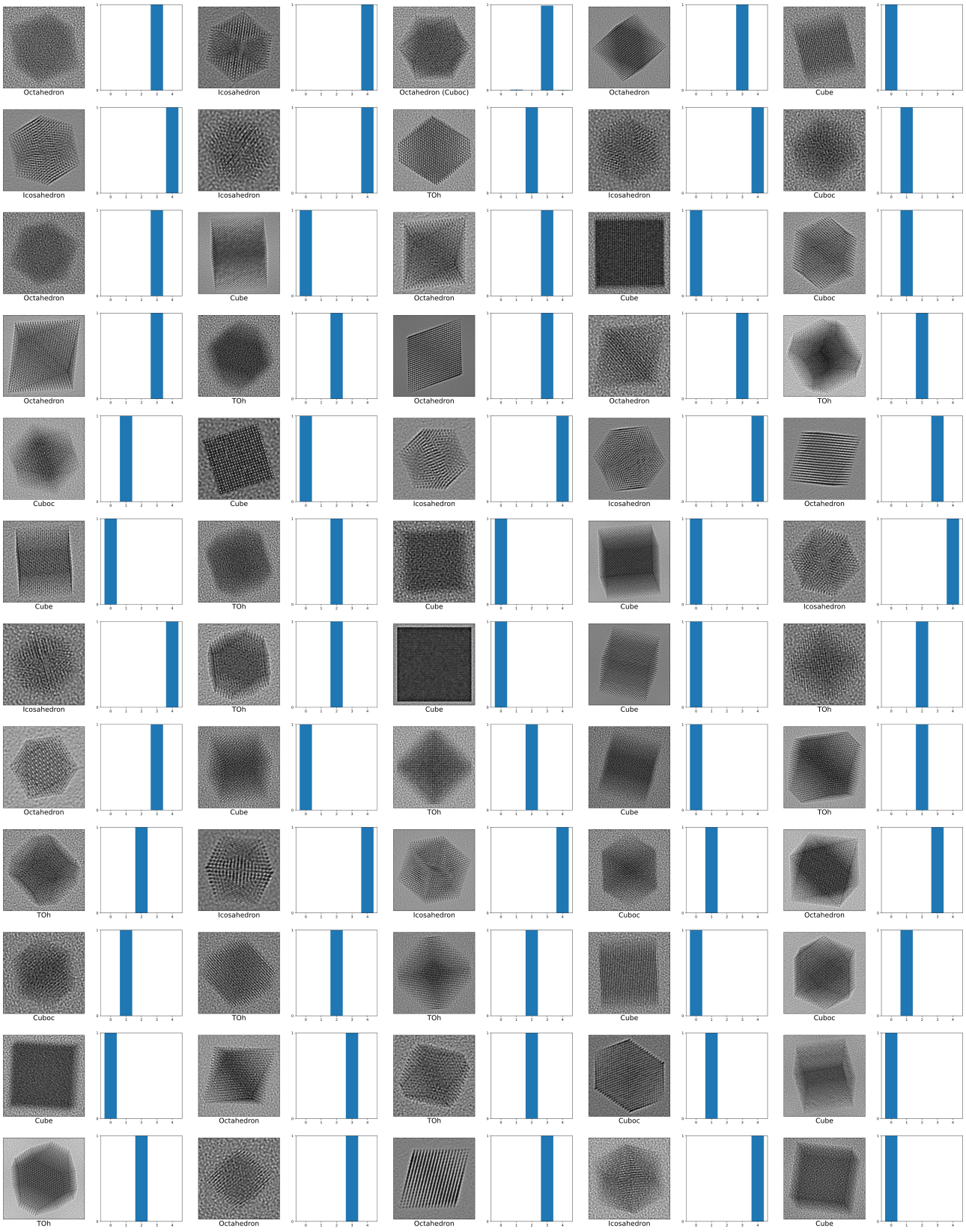

FIG. S3. 60 images were selected from a sample of 67,200 images in the test dataset. Here are the results, with the plot of the output of the classifier, and the predicted shape is indicated below each image. From 0 to 4, labels correspond to "cube", "cuboctahedron", "truncated octahedron" (="TOh"), "octahedron" and "icosahedron".

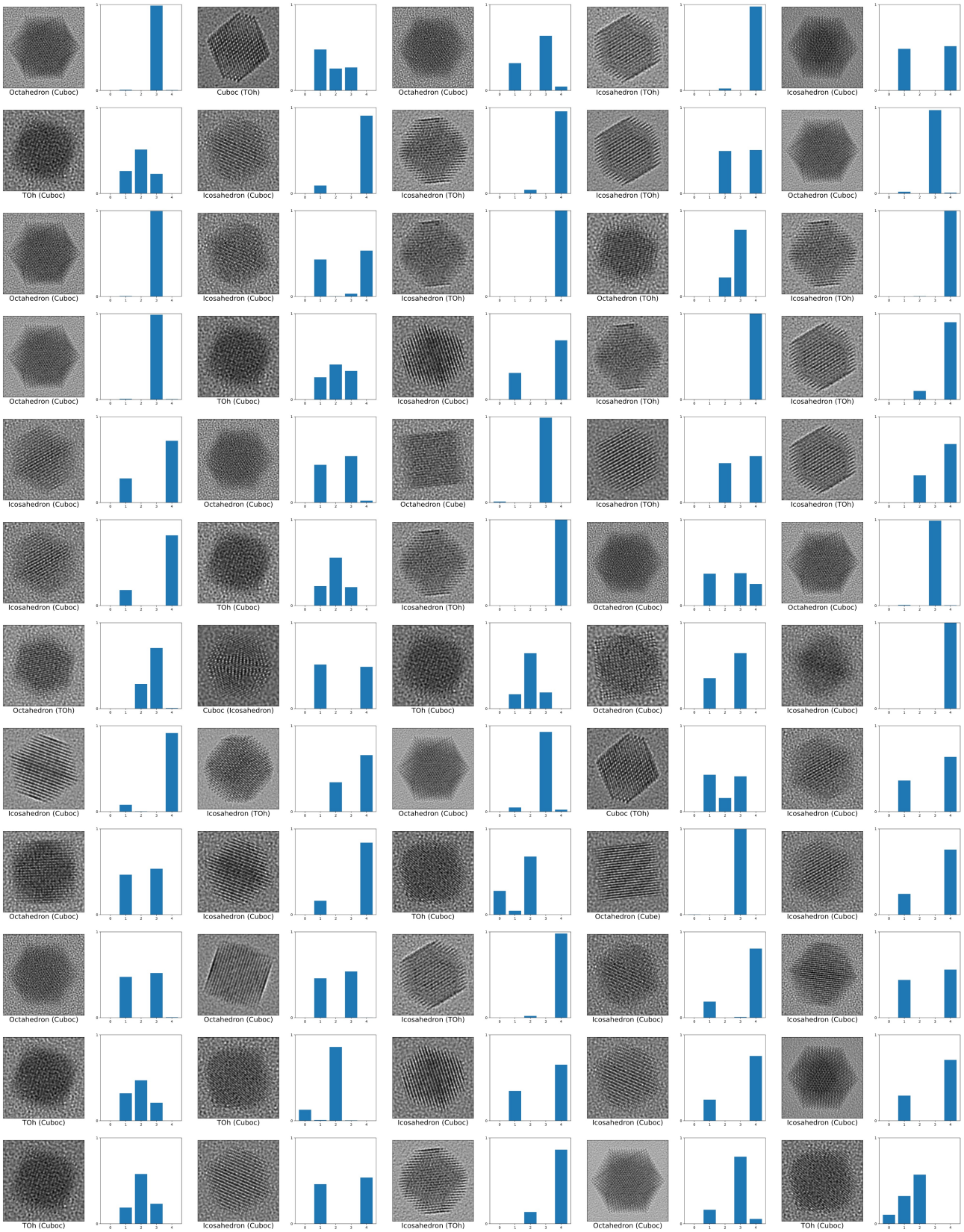

FIG. S4. 60 images were selected among less than 100 misclassifications. The predictive model was evaluated over a sample of 67, 200 images. Here are the results, with the plot of the output of the classifier, and the predicted shape is indicated below each image, along with the expected shape. From 0 to 4, labels correspond to "cube", "cuboctahedron", "truncated octahedron" (="TOh"), "octahedron" and "icosahedron".