Isotropic Reconstruction of Electron Tomograms with Deep Learning

Yun-Tao Liu[1,2,3,8], Heng Zhang[1,4,8], Hui Wang[2,3,5], Chang-Lu Tao[1,6], Guo-Qiang Bi[1,6,7,*] and Z. Hong Zhou[2,3,5,*]

[1]Center for Integrative Imaging, Hefei National Laboratory for Physical Sciences at the Microscale, and School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei, Anhui 230026, China;

[2]California NanoSystems Institute, University of California, Los Angeles, CA, USA;

[3]Department of Microbiology, Immunology and Molecular Genetics, University of California, Los Angeles (UCLA), Los Angeles, CA, USA;

[4]Department of Physics, University of Science and Technology of China, Hefei, Anhui 230026, China

[5]Department of Bioengineering, UCLA, Los Angeles, USA;

[6]Interdisciplinary Center for Brain Information, Brain Cognition and Brain Disease Institute, Shenzhen-Hong Kong Institute of Brain Science-Shenzhen Fundamental Research Institutions, Shenzhen Institute of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

[7]Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai 200031, China

[8]These authors contributed equally: Yun-Tao Liu, Heng Zhang.

*Correspondence should be addressed to:

Z.H.Z: Email: Hong.Zhou@UCLA.edu; phone: 1-310-694-7527

Or G.-Q. B: gqbi@ustc.edu.cn; +86-551-63602891

Abstract

Cryogenic electron tomography (cryoET) allows visualization of cellular structures *in situ*. However, anisotropic resolution arising from the intrinsic "missing-wedge" problem has presented major challenges in visualization and interpretation of tomograms. Here, we have developed IsoNet, a deep learning-based software package that iteratively reconstructs the missing-wedge information and increases signal-to-noise ratio, using the knowledge learned from raw tomograms. Without the need for sub-tomogram averaging, Isonet generates tomograms with significantly reduced resolution anisotropy. Applications of IsoNet to three representative types of cryoET data demonstrate greatly improved structural interpretability: resolving lattice defects in immature HIV particles, establishing architecture of the paraflagellar rod in Eukaryotic flagella, and identifying heptagon-containing clathrin cages inside a neuronal synapse of cultured cells. Therefore, by overcoming two fundamental limitations of cryoET, IsoNet enables functional interpretation of cellular tomograms without sub-tomogram averaging. Its application to high-resolution cellular tomograms should also help identify differently oriented complexes of the same kind for near-atomic resolution sub-tomogram averaging.

Introduction

The advent of single-particle cryoEM has made it routine to determine structures of isolated macromolecular complexes at 2-4 Å resolution by averaging hundreds of thousands of particles, enabling atomic modeling. The biological functions of these complexes, however, are carried out through their interactions and often depend on their spatial arrangements within cells or sub-cellular organelles[1,2]. Examples abound, ranging from pleomorphic viruses, to cellular organelles, to large-scale cellular structures like synapses between neurons. Many viruses, notably those involved in devastating pandemics such as SARS-CoV-2, influenza viruses, and human immunodeficiency viruses (HIV), are pleomorphic in

26   the organizations of their proteins and genomes. Cellular organelles, such as axonemes

27   containing microtubule doublets surrounding a central pair[3], though largely conserved in

28   their core elements across different species, rely on their non-conserved and variable

29   attachment of peripheral components that define their characteristic species-specific

30   functions[4]. In neurons, organizations of molecules, rather than molecules alone, inside the

31   synapse might underlie synaptic plasticity that is generally regarded as the cellular basis of

32   learning and memory[5,6]. Such organizational information, or "molecular sociology",

33   unfortunately is lost in single-particle cryoEM.

34   To reveal such molecular sociology across viruses or inside cells, cryogenic electron

35   tomography (cryoET) has become the tool of choice. This technique requires collecting a

36   series images of the sample at different tilt angles, called "tilt series". Due to radiation

37   damage, limited electron dosage must be further fractionated throughout the tilt series,

38   resulting in low signal-to-noise ratio (SNR) for the cryo tomogram. Furthermore, as tilting

39   increases the effective thickness of the sample, the tilt range for cryoET is usually restricted

40   to about ±70°. The missing views at higher tilt angles result in anisotropic resolution of the

41   reconstructed 3D tomograms, with the resolution along the Z-axis the lowest (Supplementary

42   Fig. 1). In Fourier space, these missing views lead to devoid of information in two

43   continuous, opposing wedge-shaped regions, commonly referred to as the "missing-wedge",

44   along the tilt axis. This missing-wedge causes severe artifacts in 3D reconstruction of

45   cellular cryoET, manifesting as, *e.g.*, oval-shaped synaptic vesicles[7] (Supplementary Fig. 1).

46   Thus, together with the low SNR in the reconstructed tomograms, the presence of missing-

47   wedge artifacts prohibits direct interpretation of the reconstructed densities in 3D, which is

48   key to the promise of cryoET to resolve molecule organization *in situ*.

49   Previous attempts have been made to partially recover information in the missing-

50   wedge[8-10] with *a priori* assumptions (*e.g.*, density positivity and solvent flatness) to constraint

51   the structural features in reconstructed tomograms. However, such assumptions have limited

52   information content (or "entropy") and may not always hold true, given the complexity of

53   biological systems. Alternatively, dual-axis tomography relies on imaging the same sample

54   with two perpendicular tilt axes, reducing the two missing-wedges to two missing pyramids;

55   thus it has the potential to alleviate artifacts in resulting tomograms[11]. However, acquisition

56   and alignment of dual-axis tilt series are substantially more complicated than that of single-

57   axis tilt series and could waste the already limited electron dose used for tilt series

58   aquisition[12]. Consequently, dual-axis tomography, while implemented in high-end

59   instruments such as the Thermo-Fisher Titan Krios, has not been practically attractive.

60   Indeed, to date, no structure with better than nanometer resolution was obtained from dual-

61   axis tomography.

62   Deep neural networks are known to learn relationships of complex data that are non-

63   linear or have high dimensionality. In the field of computer vision, convolutional neural

64   network (CNN) has been applied to various tasks, such as object recognition, image

65   segmentation, and classification, often achieving high performance. In cryoEM field, CNN-

66   based neural networks are applied to particle picking tasks and outperform conventional

67   methods such as the Laplacian of Gaussian approach[13]. CNN is also introduced to cryoEM

68   reconstruction to analyze heterogeneity of protein complexes with remarkable

69   performance[14]. However, whether CNN can also recover missing-wedge information in

70   cryoET has not been explored.

71   Here, we have developed a CNN-based software system, called *IsoNet*, for isotropic

72   reconstruction of electron tomogram. IsoNet trains deep CNN that iteratively restores

73   meaningful contents to compensate missing-wedge, using the information learned from the

74   original tomogram. The resolution at Z-axis reaches about 30Å resolution as measured by

75   the gold-standard Fourier shell correlation (FSC) criterion. By applying IsoNet to processing

76   tomograms representing viral, organelle, and cellular samples, we demonstrate its superior

77   performance in resolving novel structures of lattice defects in immature human immune-

78   deficiency virus capsid, the scissors-stack-network architecture of the paraflagellar rod, and

79   heptagon containing clathrin cage inside a neuronal synapse. The resulting tomograms with

80   isotropic resolution from IsoNet should help direct interpretation and segmentation of 3D

81   structure in cells and 3D picking hundreds of thousands of sub-tomogram particles for future

82   near-atomic resolution cryoET studies.

Results

Workflow of IsoNet

85   In spite of anisotropic resolution, tomograms generated by cryoET reconstruction contain

86   rich information with structural features such as plasma membranes, organelles, and protein

87   complexes. Thus, it is possible to recover the missing information by merging information

88   from similar features present in the same tomograms but at different orientations relative to

89   each other. An example of filling such missing information is through subtomogram

90   averaging, which aligns and averages structures of particles that are identified to be identical

91   but at different orientations in the tomogram. IsoNet is designed to expand this technique to

92   reconstruct missing-wedge information by training the neural network targeting the

93   subtomograms at different rotations for both regular and polymorphous structures.

94   The workflow of IsoNet contains five steps (Fig. 1a). Among them, three are major

95   and required: *Extract*, *Refine* and *Predict*; and the other two are optional: *Deconvolve CTF*

96   and *Generate Mask*. Each of these steps can be performed with one command of IsoNet in

97   Linux terminal. Among the 5 steps, *Refine* and *Predict* relies on graphical processing unit

98   (GPU) that provides superior processing power. The input of IsoNet is either from a single or

99   multiple tomograms. Based on the principle of machine learning, more tomograms will

100   generate more reliable results but takes longer processing time. In practice, the typical

101   number of tomograms for IsoNet is from one to five. The tomogram(s) can be reconstructed

102   by either weighted back projection (WBP) or iterative methods, such as simultaneous

103   iterative reconstructive technique (SIRT).

104   We implemented IsoNet in Python using Linux as the native operating system. The

105   package takes advantage of the Keras interface of well-established Tensorflow platform[15]

106   and can be run standard-alone, independent of other software packages. The package can

107   be run either through command line or through a graphical user interface (GUI) (Fig. 1b),

108   thus meeting to the needs of both seasoned and novice cryoET investigators. The GUI

109   contains three tabs to facilitate navigation. In each tab, information of the tomograms and the

110   parameters for each command can be specified. By clicking "*Deconvolve*", "*Generate Mask*",

111   "*Extract*", "*Refine*" and "*Predict*" buttons, user can execute the corresponding command. The

112   "only print command" option prints out the corresponding command for each step which can

113   be executed on other computers or submitted to computer clusters.

114   *Deconvolve CTF and Generate Mask steps.*   These two optional steps are performed

115   on the input tomograms prior to the subtomogram extraction in *Extract* step (Fig. 1a). The

116   *Deconvolve CTF* step has two purposes: to enhance low-resolution information and

117   compensate for the contrast transfer function (CTF) in the tomograms acquired at certain

118   underfocus conditions. Due to the presence of zeros in CTF, we used a Weiner filter for CTF

119   compensation, as implemented in Warp[16]. The *Generate Mask* step uses statistical methods

120   to detect "empty" areas in the tomograms (including vacuum above and below the sample

121   and those only containing ice or carbon) to be excluded from the subsequent analysis. Both

122   steps could improve performance and efficiency of neural network training.

123   *Extract step.*   This step allows randomly cropping subtomograms in the original tomograms

124   or the region-of-interest of the tomograms defined by masks. The maximum sizes of

125   subtomograms depend on the memory of graphics processing units (GPU), and 64^3 or 96^3

126   voxels are often used. The extracted subtomograms can be split into random halves to train

127   the neural network independently (Fig. 1a), allowing users to perform 3D gold-standard

128   FSC[17,18] to determine the resolution of IsoNet reconstructed tomograms over different

129   angular directions, particularly on Z-axis.

130   *Refine Step.*   Central to the IsoNet workflow is the *Refine* step, which iteratively trains

131   neural networks to perform missing-wedge correction and denoising (Fig. 1c). Training of the

132   neural network requires paired subtomograms as the "inputs" and the "targets." The "targets"

133   for IsoNet are the extracted subtomograms rotated at different orientations. In total, 20

134   different orientations are defined in IsoNet, generating 20 "target" subtomograms for each

135   extracted subtomogram (Supplementary Fig. 2). For each "target" subtomogram, the

136   missing-wedge is computationally imposed in Fourier space to generate the corresponding

137   "input" subtomograms (Fig. 1c). After generating the paired dataset, we train a neural

138   network to map the "input" to the "target", enabling the network to recover the imposed

139   missing-wedge artifacts. The neural network used in IsoNet adopts U-Net architecture[19],

140   containing an encoder path that extracts low-dimensional representation retaining essential

141   properties, a decoder path to reconstruct from the encoded representation, and skip-

142   connections between encoder and decoder to preserve high-resolution information

143   (Supplementary Fig. 3).

144   However, the "target" in the data pairs described above are not ideal subtomograms.

145   These subtomograms, though rotated, still miss information in other directions. To recover

146   that information and make "target" subtomograms resembling "ground truth", we adopt an

147   iterative approach: In the first iteration, we train the network with subtomograms generated

148   from the *Extract* step and obtain the IsoNet-predicted subtomograms. Then, the gained

149   information in the missing-wedge region in the Fourier space of the predicted subtomograms

150   was added to the original subtomograms, generating the first-iteration missing-wedge

151   corrected subtomograms (Fig. 1c). To further improve miss-wedge correction with more

152   iterations, the corrected subtomograms from the previous iteration are used for the paired

153   data generation in the next iteration because they are closer to missing-wedge-free 3D

154   volumes than the extracted original subtomograms. The trained network from the previous

155   iteration is then refined with the newly generated data pairs. Through multiple iterations, the

156   missing-wedge information is gradually added to the subtomograms (Fig. 1c and

157   Supplementary Fig. 4). Usually, after 10-20 iterations, the refinement converges when the

158   mean square error no longer decreases. The result of this *Refine* step is a trained network

159   that will be applied to the full tomograms and produce the isotropic reconstruction in the

160   *Predict* step (Fig. 1a).

161   Within the *Refine* step of IsoNet, we also implemented a denoising module based on

162   the noisier-input strategy[20,21]. When this optional denoising module is enabled in the *Refine*

163   step, in each iteration, 3D noise volumes are reconstructed by the back-projection algorithm

164   from a series of 2D images containing only Gaussian noise. Those 3D noise volumes are

165   then added to "input" subtomograms, with the "target" subtomograms staying the same. With

166   this strategy, IsoNet can be robustly trained with these noisier "input" subtomograms to

167   eliminate the added noise and improve the SNR of final isotropic reconstructions (Fig. 1c

168   and Supplementary Fig. 4).

169   *Predict step.*   This step performs missing-wedge correction by applying networks obtained

170   in *Refine* step to the tomograms of interest. This *Predict* step runs much faster than the

171   *Refine* step. The tomograms used for *Predict* step are typically (preferably because there are

172   no concerns of bias) the same or a subset of the tomograms used to train the network.

173   Nonetheless, users can in theory apply the trained network to tomograms of other similar

174   samples.

175   **Benchmarking with simulated data**

176   We first perform IsoNet reconstruction on simulated subtomograms using the public

177   available atomic models. Two scenarios have been considered: apoferritin[22] for the first test

178   because it has been widely used as a benchmarking specimen in high-resolution cryoEM

179   and ribosome[23] as the second test due to its asymmetric shape and primarily nucleic acid

180   content. For both tests, density maps are simulated from the atomic models using *molmap*

181   function in Chimera[24] and filtered to 8Å resolution (Figs. 1d and e). The simulated maps

182   were then rotated in 10 random directions and imposed missing-wedge in Fourier space,

183   resulting in simulated subtomograms with missing-wedge artifacts (leftmost columns in both

184   Figs. 1d and e).

185   As evident in both tests with simulated subtomograms, features such as alpha-

186   helices perpendicular to the Z-direction are smeared out in those simulated subtomograms

187   due to the missing-wedge artifact. IsoNet was then used to process those simulated

188   subtomograms. As expected, the missing information was recovered during this iterative

189   refinement process (Figs. 1d and e). After 7 iterations, all the alpha helices are visible and

190   identical to the ground truth structures in the first test. The cubic symmetry of apoferritin

191   gradually emerged even though we did not impose symmetry during the processing using

192   IsoNet. In the second test, the distortion in the shapes of ribosome is reduced during the

193   *Refine* step, with the major and minor grooves of the RNA become distinguishable (Fig. 1e).

194   These results indicate that IsoNet performed well with simulated round/symmetric protein

195   complex as well as asymmetric complex containing both protein and nucleic acid.

196   **Application to virus tomograms**

197   To further demonstrate the superior performance of IsoNet in real-world examples, we

198   perform the IsoNet reconstruction with the well-characterized cryoET datasets of virus-like

199   particles (VLP) of immature HIV-1, which is publicly available from the Electron Microscopy

200   Pilot Image Archive[25,26] (EMPIAR-10164).

201   After reconstructed with IsoNet, gold beads in the tomogram appear spherical (Fig.

202   2a), as they should, instead of the "X" shape due to the missing-wedge problem. Notably,

203   the top and the bottom of the VLP can now be observed in the IsoNet generated tomogram.

204   When examined in the Fourier space, the missing-wedge region on the XZ slices was filled

205   with values compared to the Fourier transform of the original tomogram devoid of the

206   information (Fig. 2a). To quantify the resolution of the filled information, we spilt the extract

207   subtomograms into two random subsets, trained two neural networks using those two

208   subsets independently, and then performed 3D FSC calculation[17]. The resolution on the XY

209   plane is higher than other planes (Fig. 2b), with the resolution along the X and Y axis

210   reaching the Nyquist resolution, showing our network preserves the high-resolution

211   information of the original tomograms. The Z-axis resolution of the isotropic resolution is

212   about 30Å (Fig. 2b), which is the lowest resolution in all directions. This result demonstrates

213   that our isotropic reconstruction can faithfully reconstruct the missing-wedge information at

214   least 30Å resolution.

215   Importantly, our isotropic 3D reconstruction shows that the quality of the structure is

216   similar across all directions, allowing biological structures to be interpreted adequately (Fig.

217   2c and Supplementary Video 1). We resolved those broken viruses, sheared along top and

218   bottom planes of the tomograms (Fig. 2c and Supplementary Video 2), indicating that the air-

219   water interfaces caused deformation of the capsid, as well-recognized in the cryoEM field[27].

220   The denatured Gag proteins, which are subunits of capsids, at the air-water interfaces are

221   mostly featureless.

222   The spherical viruses that were fully embedded in ice are made of hexagonal lattices

223   (Fig. 2c), whereas no pentagon subunit is observed, consistent with the subtomogram

224   averaging results of immature HIV particles[26]. Lattice defects are incorporated onto the

225   hexagonal lattices, making gaps between patches of the lattices (Fig. 2c). These defects and

226   slight curvature on the hexagonal lattices could enable the formation of the spherical shape

227   without pentagons. On lattice edges, small density protrusions extending from the hexagons

228   were observed (Fig. 2d), indicating the complete hexagons are not assembly units of HIV. In

229   concert with this observation, a recent study shows the Gag dimers are the basic assembly

230   units of the HIV-1 particle[28]. These protrusions could be Gag dimers and are prone to

231   structural changes during proteolytic cleavage[28]. Those 3D details on HIV lattices can only

232   be directly visualized after processed by IsoNet. Taken together, the above observations

233   demonstrate that IsoNet can effectively compensate for the missing-wedge problem for

234   relatively thin but heterogeneous structures, such as the immature HIV particles, and reach

235   about 30Å Z-axis resolution.

Application to tomograms of cellular organelles

237   Next, we tested the performance of IsoNet on resolving structures within cellular organelles

238   by processing tomograms of flagella of *Trypanosoma. Brucei* using IsoNet. The missing-

239   wedge compensated tomogram shows relatively uniform or isotropic structures, in all three

240   dimensions (Figs. 3a and b). The overall contrast is better than the original tomogram

241   partially due to the denoising of the network. One noticeable missing-wedge artifact is that it

242   is difficult to recognize the well-established 9 (outer doublets) + 2 (central-pair singlets)

243   microtubule arrangement in the cross-section view (i.e., XZ view in Fig. 3a). This

244   arrangement can be readily visible in the result generated by IsoNet (Fig. 3b). The missing-

245   wedge effect is also reflected by the broken and oval-shaped microtubules and severe

246   artifacts in XZ and YZ planes in the original tomogram reconstructed with SIRT algorithm

247   (Fig. 3a). In tomograms generated by IsoNet, the microtubules become complete and

248   circular-shaped with some visible tubulin subunits (Fig. 3b and Supplementary Fig. 5).

249   Binding to the microtubules, the arrays of outer (red arrows in Fig. 3b) and inner (blue arrows

250   in Fig. 3b) arm dynein proteins are now clearly distinguishable in the IsoNet generated

251   tomogram. And radial spokes connecting the outer doublets to the central pair can be

252   distinguished in all three orthogonal slices (Fig. 3a-b).

253   On one side of 9+2 microtubules lies paraflagellar rod (PFR). The structure of PFR is

254   obscure in the tomogram reconstructed by SIRT (Fig. 3c), which has given rise to the long-

255   lasting debate of the PFR organization[29-31]. The IsoNet generated tomograms showed a

256   much clearer picture of PFR. PFR density consists of parallelly arranged planes, and the

257   angle between those planes and the direction of the axoneme is 45°. Within these planes,

258   scissors-like densities stack upon each other, forming a scissors-stack-network (Figs. 3d).

259   This highly organized mesh structure could serve as a biological spring to assist the

260   movement of the flagella. This unique PFR structure observed here is consistent with the

261   organization resolved through tedious efforts of sub-tomogram averaging of thousands of

262   sub-tomograms[32]. The above observations demonstrate that IsoNet can compensate for the

263   missing-wedge problem for nonspherical cellular organelles, such as those in the Eukaryotic

264   flagella, and unveil structure with meticulous details without the need of sub-tomogram

265   averaging.

266   **Applications to tomograms of cells**

267   To evaluate IsoNet's performance for much larger and more complex structures in cells, we

268   applied IsoNet to tomograms of synapses in cultured hippocampal neurons[7]. Hippocampal

269   synapses are key devices in brain circuits for information processing and storage. They are

270   about 200-1000 nm in size, rich in proteins, lipid membranes, vesicles, mitochondria, and

271   other organelles[7,33,34]. These samples are thicker[7] (300-500 nm) than the above-described

272   flagella and virus samples, thus are representative of low SNR tomograms. The intrinsic

273   molecular crowding and structural complexity of the synapse also present difficulties for

274   missing-wedge correction. Arguably, synaptic cryoET tomograms are among the most

275   challenging datasets for any analysis algorithm.

276   IsoNet achieved isotropic reconstruction of the synaptic tomogram with substantially

277   higher contrast and better structural integrity (Figs. 4a, b, and Supplementary Videos 3-5).

278   Synaptic vesicles that were smeared out along the Z-axis in the original tomograms now

279   become spherical (Fig. 4c). The linker proteins between vesicles that are hardly seen in the

280   original tomograms now become visible in XZ and YZ planes (Figs. 4c). Even some

281   horizontally oriented features can be resolved. For example, large patches of membrane on

282   the top and the bottom of the synapse and the endoplasmic reticulum (ER) now appear

283   smooth and continuous in the isotropic reconstruction (Figs. 4b and e). These structural

284   integrity improvements facilitate the segmentation of the cellular structure, since the missing-

285   wedge corrected structures can be directly displayed based on their density threshold in 3D.

286   Particularly, placing the artificial spheres to represent synaptic vesicles, as in previous

287   studies[7,33], is no longer needed (Fig. 4e). As the elongation effect of microtubules in the Z-

288   axis being corrected, the protofilaments of microtubules have now become visible (Figs. 4d

289   and f). Inside synapses, numerous small black dots can be observed in the cytoplasm but

290   not in vesicular lumens. These dots represent small cytoplasmic proteins (orange arrows in

291   Fig. 4c), indicating our reconstruction preserves delicate structural features.

292   As a prominent example, tomograms from IsoNet revealed various types of clathrin

293   coats in hippocampal synapses. Clathrin-mediated endocytosis is a well-known presynaptic

294   vesicle recycling mechanism and is a critical step in synaptic transmission[35,36]. Clathrin

295   proteins are also present in postsynaptic compartment for neurotransmitter receptor

296   endocytosis, a process playing essential roles in synaptic plasticity[37]. Those clathrin proteins

297   are known to form cages that consist of pentagons and hexagons[38]. We observed structures

298   similar to clathrins cages of various sizes in the postsynaptic compartment in synaptic cryo

299   tomogram. However, due to the missing-wedge effect, the geometry of these clathrin cages

300   cannot be directly resolved *in situ* in typical cellular tomograms. We applied IsoNet software

301   to one synaptic tomogram that contains many putative clathrin cages in the postsynaptic

302   compartment (Figs. 5a and b, Supplementary Fig. 6). After the isotropic reconstruction, all

303   the pentagons and hexagons, which made up the clathrin cages, are revealed (Figs. 5c and

304   d). This contrasts with the original tomograms, where the polygons are smeared, especially

305   in XZ and YZ planes.

306   The 25 clathrin cages can be categorized into five types based on their geometry

307   (Supplementary Video 6). The most abundant type is minicoat, which is the smallest cages

308   the clathrin proteins can form[39]. Intriguingly, the largest clathrin cage contains two

309   heptagons, in addition to 8 hexagons and 14 pentagons (Fig. 5d, Supplementary Fig. 7),

310   which has not been reported in previous single-particle analysis[38,39]. This geometry of the

311   cage deviates from the common belief that a closed polyhedral protein cage contains 12

312   polygons. This heterogeneous in the Platonic cages of the clathrin arises from the specific

313   yet variable forms of clathrin triskelion interactions. Adapting those heptagons in neurons

314   could likely be strategy to scale up the size of the clathrin coats that enables accommodating

315   different sizes of vesicles. Intriguingly, we did not observe vesicles inside these clathrin

316   cages, suggesting that clathrin protein molecules may spontaneously self-assemble into

317   cages even when not involved in the endocytosis. It is important to note that the unexpected

318   heptagon containing clathrin cage would be lost in averaging-based methods because it only

319   has a single instance in the tomogram. Thus, these observations made in neurons

320   demonstrated that IsoNet enables compensating for missing-wedge for structures that are

321   highly heterogeneous, with limited copy numbers, and in the complex cellular environment.

Discussion

323   Here we have developed a deep learning-based package, IsoNet, to overcome the limitation

324   of missing-wedge problem and low SNR plaguing all current cryoET methods. IsoNet

325   embodies several measures that prevent the neural network from "inventing" molecule

326   features. First, the neural network was initialized with random numbers, and all the

327   information comes from original tomograms without prior knowledge. Second, we introduced

328   the dropout factor of 0.5 in the neural network so that with 50% of randomly picked neurons

329   remaining, the network can still reproduce the result. Third, to further prevent over-fitting, the

330   extracted subtomograms for training can be divided into random halves, and the resolution

331   estimation is based on the gold-standard 3D FSC.

332   To demonstrate its robustness, we have applied IsoNet to process three

333   representative types of cryoET data—pleomorphic virus HIV, cellular organelle axoneme

334   with PFR, and neuronal synapse—representing three levels of length and complexity. IsoNet

335   significantly improved structural interpretability in all these cases, allowing us to resolve

336   novel structures of lattice defects in immature HIV capsid, dynein subunits, and scissors-

337   stack-network architecture of the paraflagellar rod in eukaryotic flagella, and heptagon

338   containing clathrin cage inside a neuronal synapse. In the resulting tomograms, the *in situ*

339   protein features appear isotropic and have high quality that sometimes matches that

340   obtained through subtomogram averaging. For amorphous structures in the tomograms,

341   such as membranes, IsoNet allows the network to learn the feature representation from

342   many other similar structures in the tomogram and recover the missing information. Thus,

343   IsoNet expands the utility of cryoET by overcoming its inherent missing-wedge problem,

344   enabling 3D visualization of structures that are either complex as those in cells (Figs. 4 and

345   5) or are rare as those tomograms of patients tissues[40].

346   Philosophically speaking, no information can emerge from vacuum/nothing. So where

347   does IsoNet recover the missing information from? The questions touch upon the

348   fundamentals of deep learning and can be thought of as relating to non-locality of

349   information in space. That is, by learning from information scattered around in original

350   tomograms with recurring shapes of molecules, IsoNet sophistically eliminates distorted or

351   missing information. The great advantage of the IsoNet approach is that similar features

352   across different dimensions can be automatically discovered and "averaged" without human

353   intervention. Such features could be related in translation and rotation manners in the three

354   Cartesian dimensions, such as crystalline PFR subunits and axonemal microtubules and

355   dyneins (Fig. 3); they could also be related through symmetries, such as those pentagons

356   and hexagons of clathrin cages; ultimately, they could also be related biologically, such as

357   the facts that proteins are made up of only 20 amino acids and nucleic acid of four bases,

358   and both are geometrically constrained as a linear molecule. IsoNet learns their relationships

359   in the same tomogram or across multiple tomograms and reconstructs these features

360   automatically. In essence, therefore, IsoNet and sub-tomogram averaging compensate the

361   missing-wedge problem through the same principle.

362   Regardless of the details of information recovery, the substantial improvement in

363   map interpretability afforded by IsoNet now allows visualization of structures for functional

364   interpretation without the need of tedious and time-consuming sub-tomogram averaging,

365   which typically involves *a priori* feature identification and manual particle picking. Visualizing

366   such structures in cellular tomograms by IsoNet would also improve localization and

367   subsequent sub-tomogram averaging of hundreds of thousand copies of like-structures,

368   leading to *in situ* atomic resolution structures of cellular complexes in their native cellular

369   environment.

370 **Figures:**
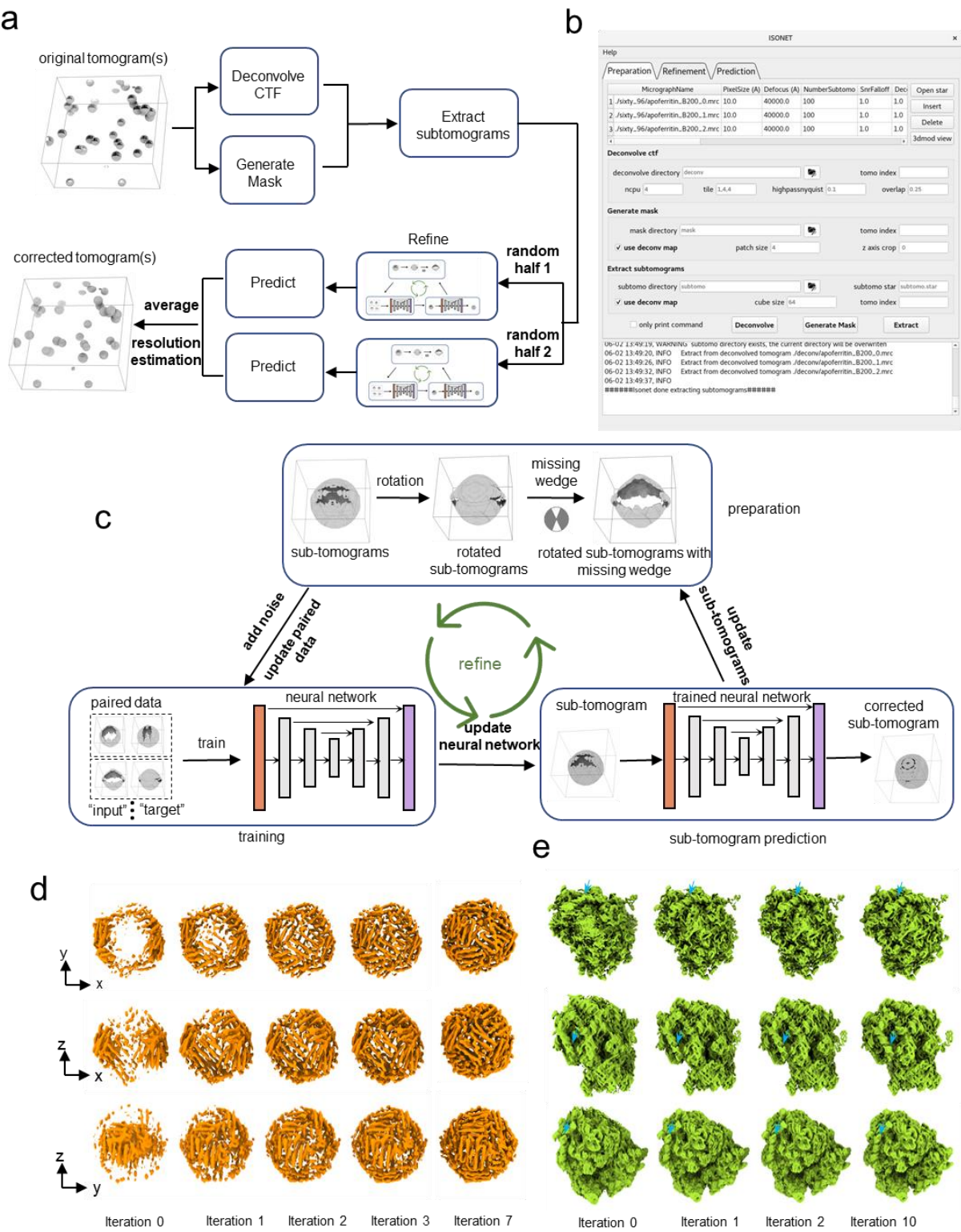
371

372 **Fig. 1 | Principle and workflow of IsoNet.**

373   **a,** Workflow of the IsoNet software. **b,** GUI of IsoNet. **c,** Illustration of *Refine* step: First,

374   subtomograms are rotated and then applied with additional missing-wedge artifacts to

375   produce paired data for training. Second, the paired data is used to train a neural network

376   with U-net architecture. Third, the trained neural network is applied to the extracted

377   subtomograms to produce missing-wedge corrected subtomograms. The recovered

378   information in these subtomograms is added to the original subtomograms, producing new

379   datasets for the next iteration. **d**, Validation of IsoNet with simulated sub-tomogram of

380   apoferritin and ribosome. Surface views from three orthogonal directions of the

381   reconstructions are shown after increasing iterations of IsoNet processing. Blue arrows

382   indicate segments of RNA duplexes.

383

384
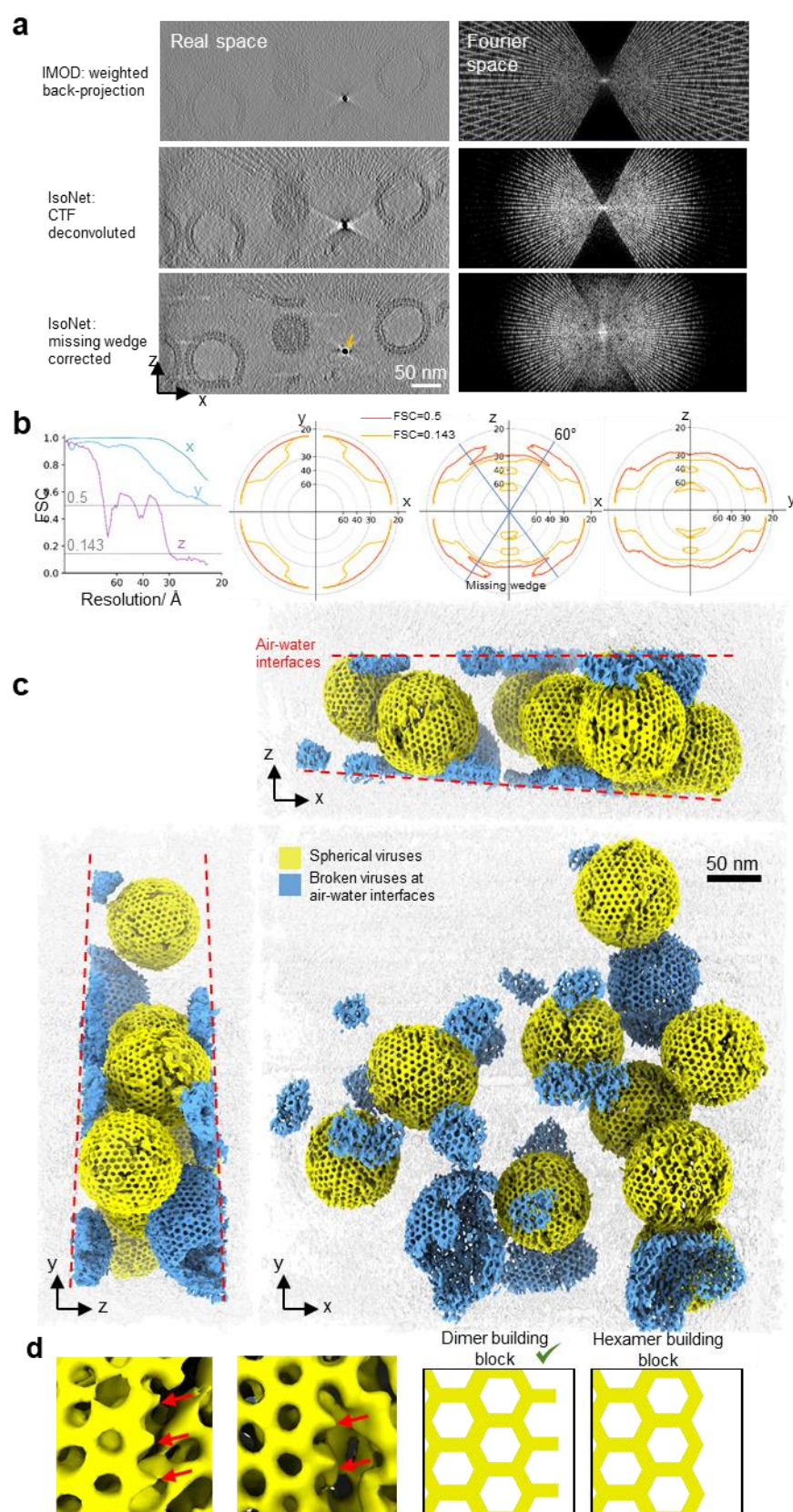
385 **Fig. 2 | IsoNet reveals lattice defects in immature HIV capsid.**

386 **a,** XZ slice views of tomogram reconstructed with WBP (top), CTF deconvoluted WBP

387 tomogram (middle), and missing-wedge corrected tomograms (bottom), with their Fourier

388 transforms on the right. Orange arrow indicates a gold bead. **b,** 3D FSC of the two

389 independent isotropic reconstructions, the left panel shows the FSC along the X, Y and Z

390 directions. Three panels on the right show the 3D FSC at 0.5 and 0.143 cutoffs on XY, XZ,

391 and YZ planes, respectively. **c,** 3D rendering of the missing-wedge corrected tomogram.

392 Dashed lines show the air-water interfaces. **d,** Examples (left) and illustrations of the lattice

393 edges of HIV capsids. Red arrows point out the density protrusions on the edges of

394 hexagonal lattices.

395

396
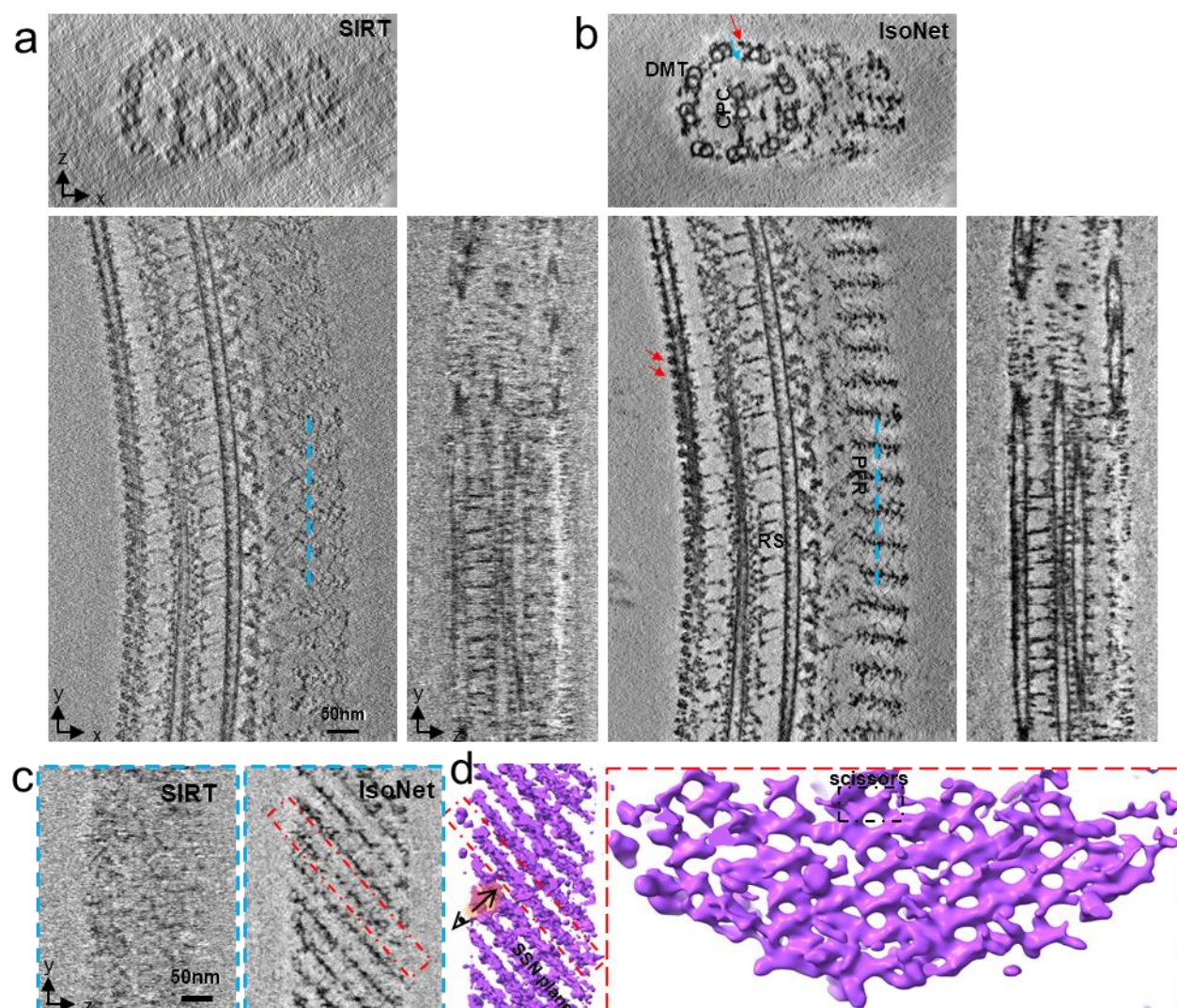
Fig. 3 | Architecture of the PFR revealed after missing-wedge correction.

a-b, Orthogonal slices of a tomogram of flagellar for SIRT reconstruction **(a)** and IsoNet

reconstruction **(b)**. DMT: Doublet microtubule; CPC: central pair complex; RS: radial spoke;

Red arrows: outer arm dyneins; Blue arrow: inner arm dynein. **c,** YZ slices show the cross-

sections corresponding to the cyan lines in **(a)** and **(b)**. **d,** 3D rendering of PFR in the IsoNet

generated tomogram. Left panel is the 3D view of PFR in the direction corresponding to **(c)**.

Right panel shows the *en face* view of a PFR scissors-stack-network (SSN) plane.

404

405

406
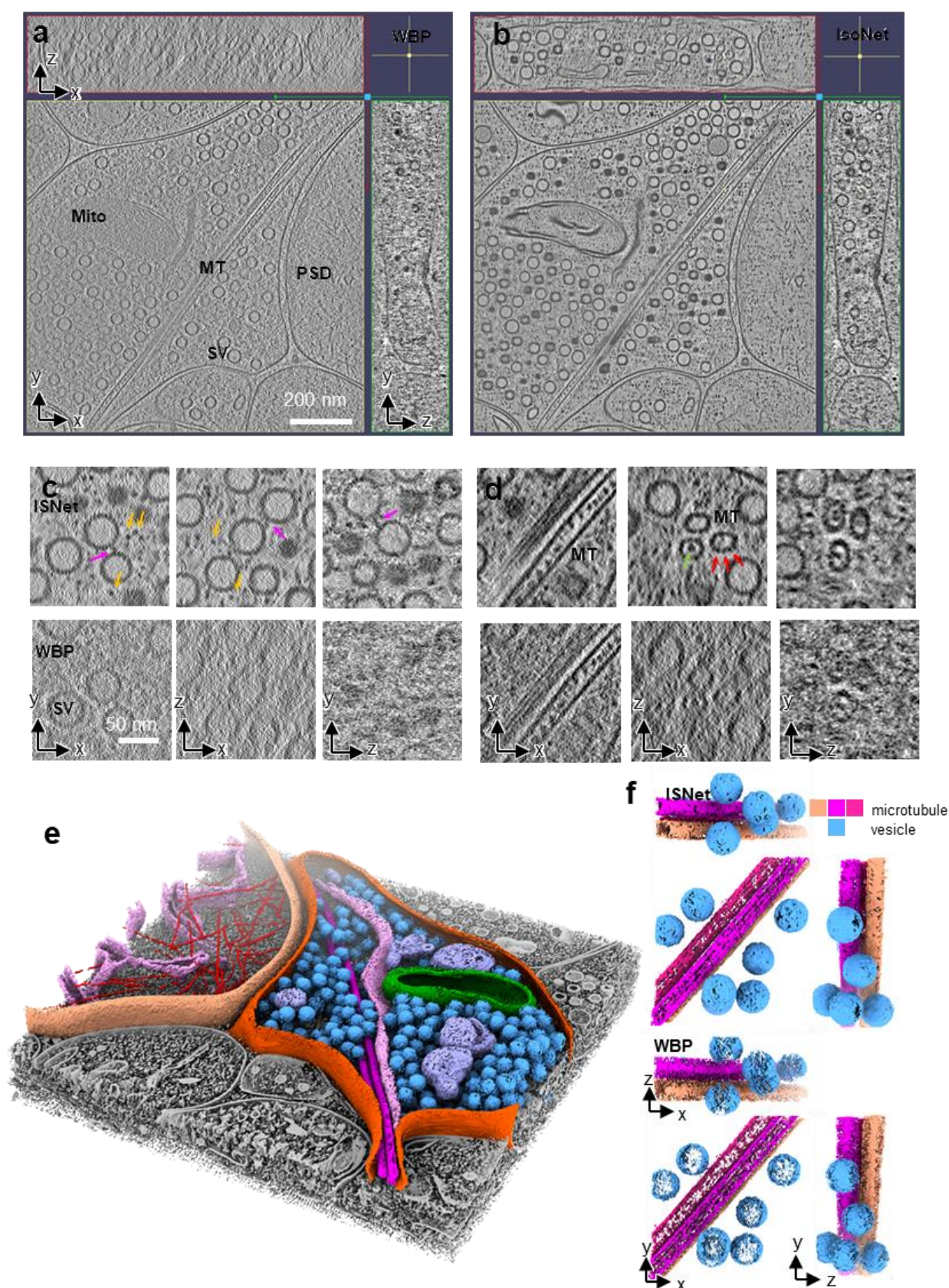
Fig. 4 | IsoNet recovers missing information in the tomograms of neuronal synapses.

408   **a-b,** Orthogonal slices of a synaptic tomogram reconstructed with WBP **(a)** and IsoNet **(b)**.

409   SV: synaptic vesicle; Mito: mitochondria; MT: microtubule; PSD: postsynaptic density. **c-d,**

410   Zoomed-in orthogonal slices of WBP reconstruction and IsoNet produced reconstruction.

411   Magenta arrows: vesicle linker; Orange arrows: small cellular proteins; Green arrow:

412   microtubule luminal particles; Red arrows: microtubule subunits. **e,** 3D rendering of the

413   tomogram shown in **(b)**. **f,** 3D rendering of a slab of tomogram with WBP reconstructions

414   and Isotropic reconstructions, showing microtubules and vesicles.

415

416
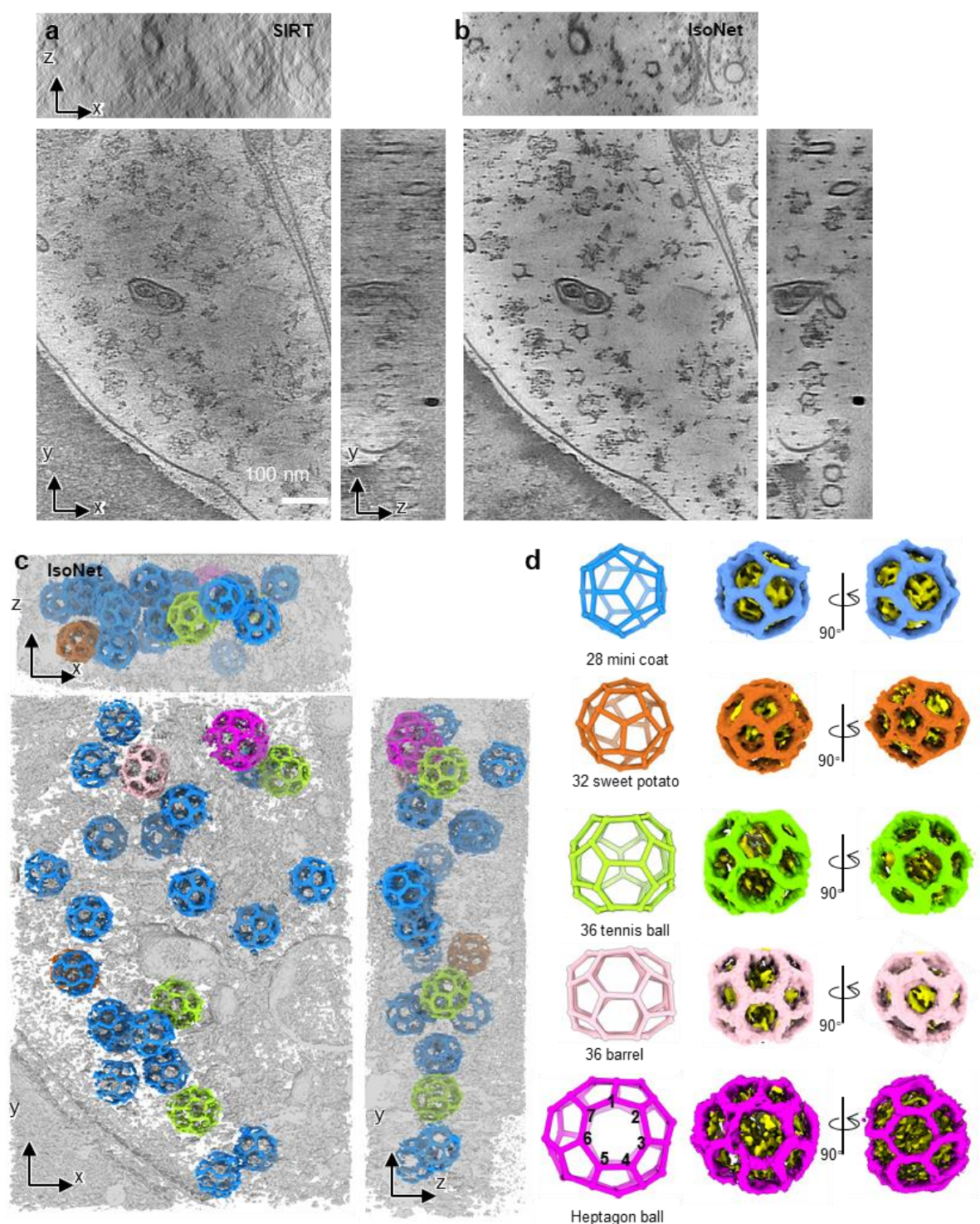
Fig. 5 | IsoNet reveals various types clathrin coats in a synapse.

a-b, Orthogonal slice views of another synaptic tomogram reconstructed with SIRT algorithm **(a)** and IsoNet **(b)**. **c,** 3D rendering of the tomogram shown in **(b)**. **d,** 3D views of the five types of clathrin cages in **(c)**.

References:

1. Beck, M. *et al.* Visual proteomics of the human pathogen Leptospira interrogans. *Nat Methods* **6**, 817-823, doi:10.1038/nmeth.1390 (2009).

2. Lucic, V., Forster, F. & Baumeister, W. Structural studies by electron tomography: from cells to molecules. *Annu Rev Biochem* **74**, 833-865, doi:10.1146/annurev.biochem.73.011303.074112 (2005).

3. Bloodgood, R. A. Sensory reception is an attribute of both primary cilia and motile cilia. *J Cell Sci* **123**, 505-509, doi:10.1242/jcs.066308 (2010).

4. Bastin, P., Pullen, T. J., Sherwin, T. & Gull, K. Protein transport and flagellum assembly dynamics revealed by analysis of the paralysed trypanosome mutant snl-1. *J Cell Sci* **112 (Pt 21)**, 3769-3777 (1999).

5. Liu, Y. T., Tao, C. L., Lau, P. M., Zhou, Z. H. & Bi, G. Q. Postsynaptic protein organization revealed by electron microscopy. *Curr Opin Struct Biol* **54**, 152-160, doi:10.1016/j.sbi.2019.02.012 (2019).

6. Liu, Y.-T. *et al.* Mesophasic organization of GABAA receptors in hippocampal inhibitory synapses. *Nature Neuroscience* **23**, 1589-1596, doi:10.1038/s41593-020-00729-w (2020).

7. Tao, C. L. *et al.* Differentiation and Characterization of Excitatory and Inhibitory Synapses by Cryo-electron Tomography and Correlative Microscopy. *J Neurosci* **38**, 1493-1510, doi:10.1523/JNEUROSCI.1548-17.2017 (2018).

8. Deng, Y. *et al.* ICON: 3D reconstruction with 'missing-information' restoration in biological electron tomography. *J Struct Biol* **195**, 100-112, doi:10.1016/j.jsb.2016.04.004 (2016).

9. Yan, R., Venkatakrishnan, S. V., Liu, J., Bouman, C. A. & Jiang, W. MBIR: A cryo-ET 3D reconstruction method that effectively minimizes missing wedge artifacts and restores missing information. *J Struct Biol* **206**, 183-192, doi:10.1016/j.jsb.2019.03.002 (2019).

10. Zhai, X. *et al.* LoTToR: An Algorithm for Missing-Wedge Correction of the Low-Tilt Tomographic 3D Reconstruction of a Single-Molecule Structure. *Sci Rep* **10**, 10489, doi:10.1038/s41598-020-66793-1 (2020).

11. Mastronarde, D. N. Dual-axis tomography: an approach with alignment methods that preserve resolution. *J Struct Biol* **120**, 343-352, doi:10.1006/jsbi.1997.3919 (1997).

12. Guesdon, A., Blestel, S., Kervrann, C. & Chretien, D. Single versus dual-axis cryo-electron tomography of microtubules assembled in vitro: limits and perspectives. *J Struct Biol* **181**, 169-178, doi:10.1016/j.jsb.2012.11.004 (2013).

13. Bepler, T. *et al.* Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs. *Nat Methods* **16**, 1153-1160, doi:10.1038/s41592-019-0575-8 (2019).

14. Zhong, E. D., Bepler, T., Berger, B. & Davis, J. H. CryoDRGN: reconstruction of heterogeneous cryo-EM structures using neural networks. *Nat Methods* **18**, 176-185, doi:10.1038/s41592-020-01049-4 (2021).

15. Abadi, M. *et al.* TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems. *arXiv*, arXiv:1603.04467 (2016).

16. Tegunov, D. & Cramer, P. Real-time cryo-electron microscopy data preprocessing with Warp. *Nat Methods* **16**, 1146-1152, doi:10.1038/s41592-019-0580-y (2019).

17. Tan, Y. Z. *et al.* Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* **14**, 793-796, doi:10.1038/nmeth.4347 (2017).

18. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat Methods* **9**, 853-854, doi:10.1038/nmeth.2115 (2012).

19. Ronneberger, O., Fischer, P. & Brox, T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Lect Notes Comput Sc* **9351**, 234-241, doi:10.1007/978-3-319-24574-4_28 (2015).

473 20 Moran, N., Schmidt, D., Zhong, Y. & Coady, P. Noisier2Noise: Learning to Denoise
474   From Unpaired Noisy Data. *CVPR*, 12064-12072 (2020).
475 21 Xu, J. *et al.* Noisy-As-Clean: Learning Unsupervised Denoising from the Corrupted
476   Image. . *arXiv* (2019).
477 22 Yip, K. M., Fischer, N., Paknia, E., Chari, A. & Stark, H. Atomic-resolution protein
478   structure determination by cryo-EM. *Nature* **587**, 157-161, doi:10.1038/s41586-020-
479   2833-4 (2020).
480 23 Zhang, X. *et al.* Structures and stabilization of kinetoplastid-specific split rRNAs
481   revealed by comparing leishmanial and human ribosomes. *Nat Commun* **7**, 13223,
482   doi:10.1038/ncomms13223 (2016).
483 24 Goddard, T. D. *et al.* UCSF ChimeraX: Meeting modern challenges in visualization
484   and analysis. *Protein Sci* **27**, 14-25, doi:10.1002/pro.3235 (2018).
485 25 Iudin, A., Korir, P. K., Salavert-Torres, J., Kleywegt, G. J. & Patwardhan, A. EMPIAR:
486   a public archive for raw electron microscopy image data. *Nat Methods* **13**, 387-388,
487   doi:10.1038/nmeth.3806 (2016).
488 26 Schur, F. K. *et al.* An atomic model of HIV-1 capsid-SP1 reveals structures regulating
489   assembly and maturation. *Science* **353**, 506-508, doi:10.1126/science.aaf9620
490   (2016).
491 27 Noble, A. J. *et al.* Routine single particle CryoEM sample and grid characterization by
492   tomography. *Elife* **7**, doi:10.7554/eLife.34257 (2018).
493 28 Tan, A., Pak, A. J., Morado, D. R., Voth, G. A. & Briggs, J. A. G. Immature HIV-1
494   assembles from Gag dimers leaving partial hexamers at lattice edges as potential
495   substrates for proteolytic maturation. *Proc Natl Acad Sci U S A* **118**,
496   doi:10.1073/pnas.2020054118 (2021).
497 29 Koyfman, A. Y. *et al.* Structure of Trypanosoma brucei flagellum accounts for its
498   bihelical motion. *Proc Natl Acad Sci U S A* **108**, 11105-11108,
499   doi:10.1073/pnas.1103634108 (2011).
500 30 Hughes, L. C., Ralston, K. S., Hill, K. L. & Zhou, Z. H. Three-dimensional structure of
501   the Trypanosome flagellum suggests that the paraflagellar rod functions as a
502   biomechanical spring. *PLoS One* **7**, e25700, doi:10.1371/journal.pone.0025700
503   (2012).
504 31 Hoog, J. L., Bouchet-Marquis, C., McIntosh, J. R., Hoenger, A. & Gull, K. Cryo-
505   electron tomography and 3-D analysis of the intact flagellum in Trypanosoma brucei.
506   *J Struct Biol* **178**, 189-198, doi:10.1016/j.jsb.2012.01.009 (2012).
507 32 Zhang, J. *et al.* Structure of the trypanosome paraflagellar rod and insights into non-
508   planar motility of eukaryotic cells. *Cell Discov* **7**, 51, doi:10.1038/s41421-021-00281-
509   2 (2021).
510 33 Fernandez-Busnadiego, R. *et al.* Quantitative analysis of the native presynaptic
511   cytomatrix by cryoelectron tomography. *J Cell Biol* **188**, 145-156,
512   doi:10.1083/jcb.200908082 (2010).
513 34 Liu, Y. T. *et al.* Mesophasic organization of GABAA receptors in hippocampal
514   inhibitory synapses. *Nat Neurosci* **23**, 1589-1596, doi:10.1038/s41593-020-00729-w
515   (2020).
516 35 Mettlen, M., Chen, P. H., Srinivasan, S., Danuser, G. & Schmid, S. L. Regulation of
517   Clathrin-Mediated Endocytosis. *Annu Rev Biochem* **87**, 871-896,
518   doi:10.1146/annurev-biochem-062917-012644 (2018).
519 36 Rodal, A. A. & Littleton, J. T. Synaptic endocytosis: illuminating the role of clathrin
520   assembly. *Curr Biol* **18**, R259-261, doi:10.1016/j.cub.2008.02.014 (2008).
521 37 Man, H. Y. *et al.* Regulation of AMPA receptor-mediated synaptic transmission by
522   clathrin-dependent receptor internalization. *Neuron* **25**, 649-662, doi:10.1016/s0896-
523   6273(00)81067-3 (2000).
524 38 Paraan, M. *et al.* The structures of natively assembled clathrin-coated vesicles. *Sci
525   Adv* **6**, eaba8397, doi:10.1126/sciadv.aba8397 (2020).

39   Morris, K. L. *et al.* Cryo-EM of multiple cage architectures reveals a universal mode of clathrin self-assembly. *Nat Struct Mol Biol* **26**, 890-898, doi:10.1038/s41594-019-0292-0 (2019).

40   Wang, R. *et al.* Electron cryotomography reveals ultrastructure alterations in platelets from patients with ovarian cancer. *Proc Natl Acad Sci U S A* **112**, 14266-14271, doi:10.1073/pnas.1518628112 (2015).

41   Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* **180**, 519-530, doi:10.1016/j.jsb.2012.09.006 (2012).

42   Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. *arXiv* (2014).

43   Zheng, S. Q. *et al.* MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods* **14**, 331-332, doi:10.1038/nmeth.4193 (2017).

44   Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* **192**, 216-221, doi:10.1016/j.jsb.2015.08.008 (2015).

45   Imhof, S. *et al.* Cryo electron tomography with volta phase plate reveals novel structural foundations of the 96-nm axonemal repeat in the pathogen Trypanosoma brucei. *Elife* **8**, doi:10.7554/eLife.52058 (2019).

46   Mastronarde, D. N. Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol* **152**, 36-51, doi:10.1016/j.jsb.2005.07.007 (2005).

47   Turonova, B., Schur, F. K. M., Wan, W. & Briggs, J. A. G. Efficient 3D-CTF correction for cryo-electron tomography using NovaCTF improves subtomogram averaging resolution to 3.4A. *J Struct Biol* **199**, 187-195, doi:10.1016/j.jsb.2017.07.007 (2017).

48   Kremer, J. R., Mastronarde, D. N. & McIntosh, J. R. Computer visualization of three-dimensional image data using IMOD. *J Struct Biol* **116**, 71-76, doi:10.1006/jsbi.1996.0013 (1996).

Methods

Software implementation

We implemented IsoNet in Python using Linux as the native operating system. Typical hardware setup includes one node with four Nvidia GeForce 1080Ti GPU cards of 11 gigabytes memory, which is common in a cryoET research laboratory. The package can be run from command line and relies on Keras that acts as interface for Tensorflow[15], and the package can be downloaded from Github (https://github.com/Heng-Z/IsoNet). A detailed document is provided, accompanied by the IsoNet software. Tutorial dataset and video can be found in: https://github.com/Heng-Z/IsoNet_tutorial.

This package is standard-alone and does not rely on other software such as IMOD, while some common Python modules must be installed prior to running IsoNet. Such Python modules are easy to install with the "pip" command. For example, the IsoNet uses Python module "mrcfile" for the read and write tomogram or subtomogram, and "numpy" for the image processing such as rotation and Fourier transform. The U-net neural network is built by stacking multiple layers (Supplementary Fig. 3) provided in "tensorflow.keras.layers". For example, three "Conv3D" layers are stacked together in each depth of the encoding path of the U-net.

The package can be launched through a single command entry, either "*isonet.py*" or "*isonet.py gui*", for Linux command line operations or a graphical user interface (GUI) (Fig. 1b), respectively. Users can then access all the processing steps of the IsoNet procedure. The IsoNet procedure contains five steps, including three major steps: Extract, Refine, and Predict and two accessary steps: CTF deconvolve and mask generate. Each of these steps corresponds to one command of IsoNet in Linux terminal and will be described in the following sections.

Dataset preparation

To use IsoNet, users should prepare a folder containing all tomograms. Binning the tomograms to a pixel size of about 10Å is recommended. Typically, a folder containing 1 to 5 tomograms is used as input. These input tomograms can either be reconstructed by SIRT or WBP algorithm. The tilt axis is the Y-axis, and recommended tilt range is -60 to 60°, while other tilt ranges might also work. The tilt series can be collected with any tilt schemes, continuous, bidirectional, or dose-symmetric.

587   IsoNet uses Self-defining Text Archive and Retrieval (STAR) file format to store

588   information of tomograms and program parameters been used during Isonet processing.

589   Thus, it inter-operates seamlessly with such leading cryoEM software packages as Relion[41].

590   Tomogram STAR file, default named as *tomograms.star*, is required to run IsoNet and this

591   file can be prepared with IsoNet GUI, with a text editor, or with the following command:

592                            $ *isonet.py prepare_star [tomogram_folder]*

Deconvolve CTF

594   For the tilt series imaged without Volta phase plate (VPP), the sinusoidal CTF suppresses or

595   even inverts information at certain frequencies. To enhance the contrast of the tomograms

596   and promote information retrieval, CTF deconvolution, similar to what is implemented in

597   Warp software, is applied to tomograms in this step.

598   IsoNet uses Wiener-like filter[16] for CTF deconvolution, with spectral signal-to-noise

599   ratio (SSNR) set empirically:

600
$$SSNR = e^{-f \times 100F} \times 10^S \times H_f$$

601   Where f denotes the spatial frequency, H a high-pass filter, F the custom fall-off

602   parameter, and S denotes the custom strength parameter. Because the SSNR of the

603   Weiner-like filter is determined empirically, users can tune the SSNR fall-off or deconvolve

604   strength parameters to enhance contrast of the tomograms. This step can be performed by

605   "CTF deconvolve" function in GUI or with the following command:

606            $ *isonet.py deconv [tomogram_star] --snrfalloff 1.0 --deconvstrength 1.0*

Generate mask

608   Subtomograms for training would be better to contain rich information than empty areas with

609   only ice, air, or carbon. In this optional mask generation step, IsoNet uses statistical methods

610   to detect empty regions from which subtomograms will not be extracted. Two different types

611   of masks can be applied: density mask that excludes areas with low cryoET density and

612   standard deviation mask that excludes areas with low standard deviation.

613   The density mask will first suppress the noise with a Gaussian filter and then apply a

614   sliding window maximum filter to the contrast-inverted tomogram. The areas with relatively

615   smaller density values in the filtered tomogram will be deemed as empty space and

616 excluded in the mask. The parameter "*density_percentage*" defines the percentage of the
617 area kept in the tomogram by the density mask.

618 The standard deviation mask is achieved by calculating the standard deviation of
619 voxels in a small cubic volume centered at each evaluating voxel. The voxels having
620 relatively lower standard deviation will be excluded. The parameter "*std_percentage*" defines
621 the percentage of voxels kept by the standard deviation mask.

622 IsoNet uses the intersection of these two masks. The parameters for mask
623 generation can be tuned to cover the region of interest but exclude empty areas. In addition
624 to these two types of masks, IsoNet allows excluding top and bottom parts of tomograms
625 which usually are empty areas by the "*z_crop*" parameter.

626 Usually, the default parameters will provide a good mask to exclude empty areas;
627 users can also tune the parameters using GUI by "Generate mask" function or the following
628 command, for example:

629 $ *isonet.py make_mask [tomogram_star] --density_percentage 50*

630 **Extract subtomograms**

631 In each tomogram, specified number of seeds are randomly generated within the whole
632 tomogram or the region of interest defined by the mask. Then, cubic volumes centered at the
633 generated seeds are boxed out and saved as subtomograms. The extracted subtomograms
634 should be large enough to cover typical features in tomograms, such as a patch of
635 membrane or vesicle. However, due to the GPU memory limitation, this size cannot be
636 arbitrarily large. We usually extract 300 subtomograms of 96^3 voxels in total. After extraction,
637 the contrast of those subtomograms is inverted. Then, tomograms are normalized by
638 percentile to ensure 80% of the voxel values fit into the range between zero and one. The
639 subtomograms can be randomly split into two halves and used for performing missing-
640 wedge correction independently to eliminate overfitting and calculate gold-standard FSC.

641 The information of extracted subtomograms is stored in another STAR file, default
642 named as *subtomo.star*. Subtomogram extraction can be performed by either IsoNet GUI or
643 the following command:

644 $ *isonet.py extract [tomogram_star]*

645 **Refine**

646 This process iteratively trains neural networks to fill the missing-wedge information using the

647 same tomograms whose missing-wedge artifacts were added to other directions. The

648 denoising module can also be enabled in this step, making the network capable of reducing

649 noise and recovering the missing-wedge. After refinement, the resulting subtomograms and

650 neural network model in each iteration are saved. The network models with a suffix of ".h5"

651 can be used for the prediction step.

652 Four steps, including *training dataset generation*, *adding noise*, *network training*, and

653 *subtomograms prediction*, will be performed during each iteration. These steps will be

654 described in the following sections. The missing-wedge restored subtomograms by

655 *subtomograms prediction* in every iteration will be used for *training dataset generation* in the

656 next iteration. Usually, 10-15 iterations in the refine step are sufficient to obtain a well-trained

657 network for the missing-wedge correction, whereas more iterations can be performed for

658 refinement with denoising.

659 The refine step can be performed from the GUI or with the following command, for

660 example:

661 $ *isonet.py refine [tomogram_star] --iterations 30 --gpuID '0,1,2,3'*

662 Users can also continue training from the previous iterations using "*continue_from*" or

663 from previously trained models using "*pretrained_model*" parameter.

Refine step 1: training dataset generation

665 To generate paired datasets for neural network training, IsoNet rotates the extracted

666 subtomograms to different orientations. Twenty rotated copies can be obtained for each

667 extracted subtomogram as follows (Supplementary Fig. 2). First, each subtomogram is a

668 cube with six faces. Each face can be rotated with an out-of-plane angle to face toward the

669 positive direction of the Z-axis. Second, each out-of-plane rotation can be followed by four

670 in-plane rotations, making 24 possible rotations. However, among the 24 rotations, four of

671 them result in subtomograms with the same missing-wedge direction as the original

672 subtomograms. Thus, these four rotations are excluded, resulting in 20 orientations for each

673 subtomograms. This rotation process enlarges the original dataset by 20 times for training,

674 making it possible to achieve a good performance of missing-wedge correction even with a

675 small dataset, *e.g.,* a single tomogram.

676   After the rotation, the IsoNet program then applies missing-wedge filter to the rotated

677   subtomograms. The missing-wedge filter volume has the same size as that of the

678   subtomograms. In the missing-wedge filter volume, voxel value is zero inside the missing-

679   wedge region and one in the rest of the volume. Then, the Fourier transforms of the rotated

680   subtomograms are multiplied by the missing-wedge filter volume and then transformed back

681   to the real space, generating missing-wedge filtered subtomograms.

682   To avoid the incomplete information along the edge of the subtomograms when

683   applying missing-wedge filter, both rotated subtomograms and missing-wedge filtered

684   rotated subtomograms are trimmed into smaller volumes (often 64^3 voxels), generating

685   "target" and "input" for the network training, respectively. These generated data pairs are

686   used to train neural network that maps the "input" to "target".

687   **Refine step 2: adding noise.**

688   This optional step allows performing missing-wedge correction and denoising simultaneously

689   using IsoNet. IsoNet uses a noisier-input strategy[20,21] that learns to map "input" with

690   additional noise to the "target".

691   IsoNet simulates the noise pattern in reconstructed tomograms with the assumption

692   that in every acquired projection, the noise is an additive Gaussian noise and independent

693   among all images acquired in a tilt series. During the adding noise step, a set of 3D noise

694   volumes are constructed by back-projecting a series of 2D Gaussian noise images to reflect

695   the effect of the back-projection algorithm on noise formation.

696   The denoise level is defined as the ratio of the standard deviation between the

697   subtomograms and the added noise. The noise volumes are scaled to match the denoise

698   level before being added to the "input" subtomograms. Thus, the lower denoise value means

699   less noise is added to individual subtomograms.

700   Because the added noise may further corrupt the 3D subtomograms, making the

701   network hard to train, it is recommended to start the first several iterations of refinement

702   without denoising. After the refinement results are nearly converged, the noise volume can

703   then be added to the "input" subtomograms in the following iterations. A typical routine is to

704   train ten iterations without denoising and then increase the denoise level by 0.05 for every

705   five iterations. This step-wised noise addition can be performed automatically in the refine

706   step of the IsoNet software.

Refine step 3: network training

Neural network used in IsoNet is based on U-net, which is well recognized in biomedical image restoration and segmentation[19]. The main building blocks of the U-net are 3D convolution layers with non-linear activation functions called Rectified Linear Units (Relu), which are applied per voxel. Those convolution layers have kernel sizes of 3x3x3. Three 3D convolution layers are stacked together to form a convolution block in our network, which can extract complicated features.

By stacking the convolutional blocks, the U-net is built based on encoder-decoder architecture (Supplementary Fig. 3). The encoder path is a set of convolution blocks and strided convolution layers that compress 3D volumes. Strided convolution layers reduce the spatial size of the input of this layer by 2x2x2, allowing the network to learn more abstract information. A convolution block followed by a strided convolution layer makes one encoder block in the contracting path. Total three encoder blocks form the entire encoding path. The number of convolution kernels for each convolution layer doubles after each encoder block. After the encoder path, the 3D volumes are processed with a convolution block and enter the decoder path of the network. The decoder path is symmetrical to the encoder but uses transpose convolution layers, opposite to strided convolution layers, to enlarge the dimension of features.

Although the down-sampling of the 3D volumes captures the essence of the features, high-resolution information is lost by stride convolution operations. In particular, the skip connections that concatenate the feature layers of the same dimension in two paths are implemented to preserve the high-resolution information. Dropout strategy that randomly set 50% of neurons' activation to 0 in the convolution layers is used to prevent overfitting during the training.

This network uses the mean absolute error between the output of the network and the target subtomograms as loss function. The loss function is minimized by employing Adam optimizer[42] with an initial learning rate of 0.0004. The neural network training is performed on GPU and consists of several epochs or cycles (typically ten epochs). Each epoch will traverse through the paired dataset. The data pairs are grouped into batches (which generally have a size of 8 or 16) to feed into each epoch. After the training, the trained neural networks are saved for the next iteration of the refine step.

Refine step 4: Subtomogram prediction

739   After each iteration of refinement, the network is applied to the original subtomograms,

740   generated predicted subtomograms. Then IsoNet generates an inverse missing-wedge filter

741   volume with value one inside the missing-wedge region but zero in the rest of the 3D

742   volume. The predicted subtomograms are then transformed to Fourier space and multiplied

743   with the inverse missing-wedge filter volume to extract the added information inside the

744   missing-wedge regions. Then, these filtered volumes are added to the original

745   subtomograms, generating the missing-wedge restored subtomograms for subsequent

746   iterations of refinement.

747   **Predict**

748   After the refine step, the trained network is saved in a model file. It will be used to correct the

749   missing-wedge for the original tomogram or other similar tomograms. For most tomograms,

750   the full-size 3D images can hardly fit into the memory of a regular GPU. Thus, the IsoNet

751   program splits the entire tomogram into smaller 3D chunks to apply the network on them

752   separately. Then output 3D chunks are montaged to produce the final output. To avoid the

753   line artifact between adjacent chunks caused by the loss of information on the edges of

754   subtomograms. We implemented a seamless reconstruction method called overlap-tile

755   strategy, which predicts the overlapping chunks to avoid the edge effect. The "crop_size"

756   parameter defines the size of the cubic chunks. This predicting step can be performed with

757   IsoNet GUI or with the following command, for example:

758         $*isonet.py predict [tomogram_star] [output_folder] --model [network_model]*

759   **Benchmarking with simulated data**

760   We performed IsoNet reconstruction on simulated subtomograms using the public available

761   atomic models: apoferritin model[22] (PDB: 6Z6U) and ribosome model[23] (PDB: 5T2C). For

762   both tests, density maps were simulated from the atomic models using "*molmap*" function in

763   ChimeraX[24] and filtered to 8Å resolution (Fig. 1d, e). Those simulated maps with 2.67 Å/pix

764   pixel size were then rotated in 10 random directions and imposed with missing-wedge filter

765   in Fourier space, resulting in simulated subtomograms with missing-wedge artifacts (leftmost

766   columns in both Fig. 1d, e).

767   For the simulated Apoferritin maps, we created a subtomogram STAR file with the

768   "*isonet.py prepare_subtomo_star*" command. With this subtomogram STAR file as input, we

769   performed an IsoNet refine step for ten iterations without denoising. For benchmarking with

770   the simulated ribosome maps, we extracted eight smaller subtomograms (96^3 voxels) from

771 each ribosome map due to the larger dimension of the ribosome map. The subtomogram

772 STAR file generated in the extract step was used for subsequent refine step. After ten

773 iterations, trained network was obtained and was then used to produce missing-wedge

774 corrected maps of ribosome using "isonet.py predict" command.

Processing tomograms of HIV virus

776 For pleomorphic viruses, we downloaded an HIV dataset from public repository EMPIAR-

777 10164[26]. Three tilt series, TS_01, TS_43 and TS_45, were used for testing. The movie stacks

778 were drift corrected with MotionCorr[43] and reconstructed with IMOD[11] using WBP algorithm.

779 The defocus value of each image was determined by CTFFIND4[44]. Eight-time binned

780 tomograms with 10.8Å pixel size were used for further processing. For the CTF

781 deconvolution of the tomograms, SSNR fall-off and the deconvolve strength parameters

782 were set to 0.7 and 1, respectively. Then, we created one mask for each tomogram using

783 "isonet.py make_mask" command. Total 300 subtomograms with 96^3 voxels were randomly

784 extracted from the three tomograms and then split into random halves. For each half of the

785 subtomograms, we performed refine step for 35 iterations independently, generating two

786 trained neural networks. In the predict step of IsoNet, tomogram TS_01 was used to

787 generate two missing-wedge corrected tomograms by the two independently trained

788 networks. These two tomograms were then averaged to create a final map.

789 These two missing-wedge corrected tomograms enabled calculating gold-standard

790 FSC. Instead of calculating a global FSC, we performed 3D FSC calculation for all the

791 directions[17] to measure the resolution anisotropy of the missing-wedge corrected tomogram.

792 Because the 3D FSC calculation works for cubic volumes while the size of the tomogram is

793 non-cubical, we cropped the generated tomograms into cubic subtomograms for the 3D FSC

794 calculation. As for the HIV dataset, the 3D FSC was calculated for four 200^3 volumes

795 cropped from both missing-wedge corrected HIV tomograms. The resulting four 3D FSC

796 were then averaged to produce the final 3D FSC, the orthogonal sections of which are

797 shown in Figure 2b.

Processing tomograms of the Eukaryotic flagella

799 For cellular organelles, we chose the demembraned flagella of *Trypanosoma. Brucei.* The

800 datasets described here were obtained in our previous studies[32,45]. Tilt series were recorded

801 with SerialEM[46] by tilting the specimen stage from −60° to +60° with 2° increments. The

802 cumulative electron dosage was limited to 100 to 110 e⁻/Å² per tilt series. The movie stacks

803   were drift corrected with MotionCorr[43] and reconstructed with IMOD using SIRT algorithm.

804   The tomograms were binned by four, resulting in a pixel size of 10.21 Å/pixel.

805   Three tomograms were chosen for missing-wedge correction. These tilt series were

806   acquired with VPP so that we did not perform the CTF deconvolution. We generated one

807   mask for each tomogram using "isonet.py make_mask" command. Then, we extracted a

808   total of 360 cubic subtomograms with 128^3 voxels from three tomograms. Using these

809   subtomograms, we trained a network by running the refine step for 35 iterations with default

810   denoise levels, which were automatically changed from 0 to 0.2. Trained network produced

811   in the refine step was then used to run the predict step of IsoNet to obtain a final missing-

812   wedge corrected tomogram, which is shown in Figure 3.

813   **Processing tomograms of hippocampal neurons**

814   Tomograms of hippocampal neurons were obtained in our previous study[7]. The two

815   tomograms used in this study were collected on a Titan Krios microscope equipped with K2

816   summit in counting mode. The energy filter (Gatan image filter) slit was set at 20 eV. The

817   Titan Krios was operated at an acceleration voltage of 300 KV. Tilt series were acquired

818   using SerialEM[46] with tilt scheme: from +48° to -60° and from +50° to +66° at an interval of

819   2°. The total accumulated dose was ~150 e-/Å^2. The pixel size of the tomograms is 4.35

820   Å/pixel. Each recorded movie stack was drift-corrected and averaged to produce a

821   corresponding micrograph using MotionCorr[43]. The tilt series were aligned using IMOD[11].

822   One tilt series of the tomogram shown in Figure 4 was imaged with VPP, while the other

823   shown in Figure 5 was acquired without VPP. When VPP was used, the defocus value was

824   maintained at -1 μm; otherwise, it was kept at -4 μm.

825   For the tilt series recorded with VPP, the aligned tilt series were reconstructed using

826   NovaCTF[47], generating a tomogram reconstructed with WBP. The tomogram was binned by

827   four, and 300 subtomograms (96^3 voxels) were extracted from that tomogram. Those

828   subtomograms were then used 35 iterations of refinement. The trained network produced in

829   refine step was used for missing-wedge correction for the entire tomogram (Fig. 4).

830   For the tilt series recorded without VPP, the defocus value of each image was

831   determined by CTFFIND4[44], and CTF phase flipped tomogram was obtained by NovaCTF[47].

832   This tomogram (Fig. 5) was reconstructed with SIRT-like filter, with CTF phase flipping

833   performed on the individual tilt images. The tomogram was binned by four for missing-wedge

834   correction with IsoNet. Then, 200 subtomograms (96^3 voxels) were extracted from the

835   tomogram in the extract step of IsoNet. A trained network was obtained with the refine step

836   of IsoNet for 35 iterations. The trained network was then used for the predict step of IsoNet,

837   producing a missing-wedge corrected tomogram (Fig. 5).

838   **3D visualization**

839   IMOD[48] was used to visualize the 2D tomographic slices. UCSF ChimeraX[24] was used to

840   visualize the resulting IsoNet generated tomograms in their three dimensions. Segmentation

841   of densities maps and surface rendering were performed by the volume tracer and color

842   zone in UCSF ChimeraX.

843   # End notes

844   **Supplementary information.** Supplementary Information is linked to the online version of

845   the paper.

857

858   **Author contributions:** Y.-T.L. conceptualized the method, wrote code, processed data,

859   made illustrations and documentation, and wrote the paper; H.Z. wrote code, processed the

860   HIV and neuronal synapse data, made documentation and assisted in illustrations, and

861   wrote the paper; H. W. participated in coding, processed the axoneme with PFR data, made

862   documentation and assisted in illustrations; C.-L.T. assisted in testing the method with the

863   neuronal synapse data and in making illustration based on the test results; G.-Q. B. and

864   Z.H.Z. oversaw the project, interpreted the results, and wrote the paper. All the authors

865   edited and approved the manuscript.

866   **Author information:** The authors declare no competing financial interests. Correspondence

867   and requests for materials should be addressed to G.-Q.B. (gqbi@ustc.edu.cn) or Z.H.Z.

868   (Hong.Zhou@UCLA.edu).

869

870   **Data and materials availability:** All data reported in the main text or the supplementary

871   materials are available upon request. The software has been deposited to GitHub page at

872   https://github.com/Heng-Z/IsoNet.

873

874 **Supplementary Figures and Videos:**

875

876
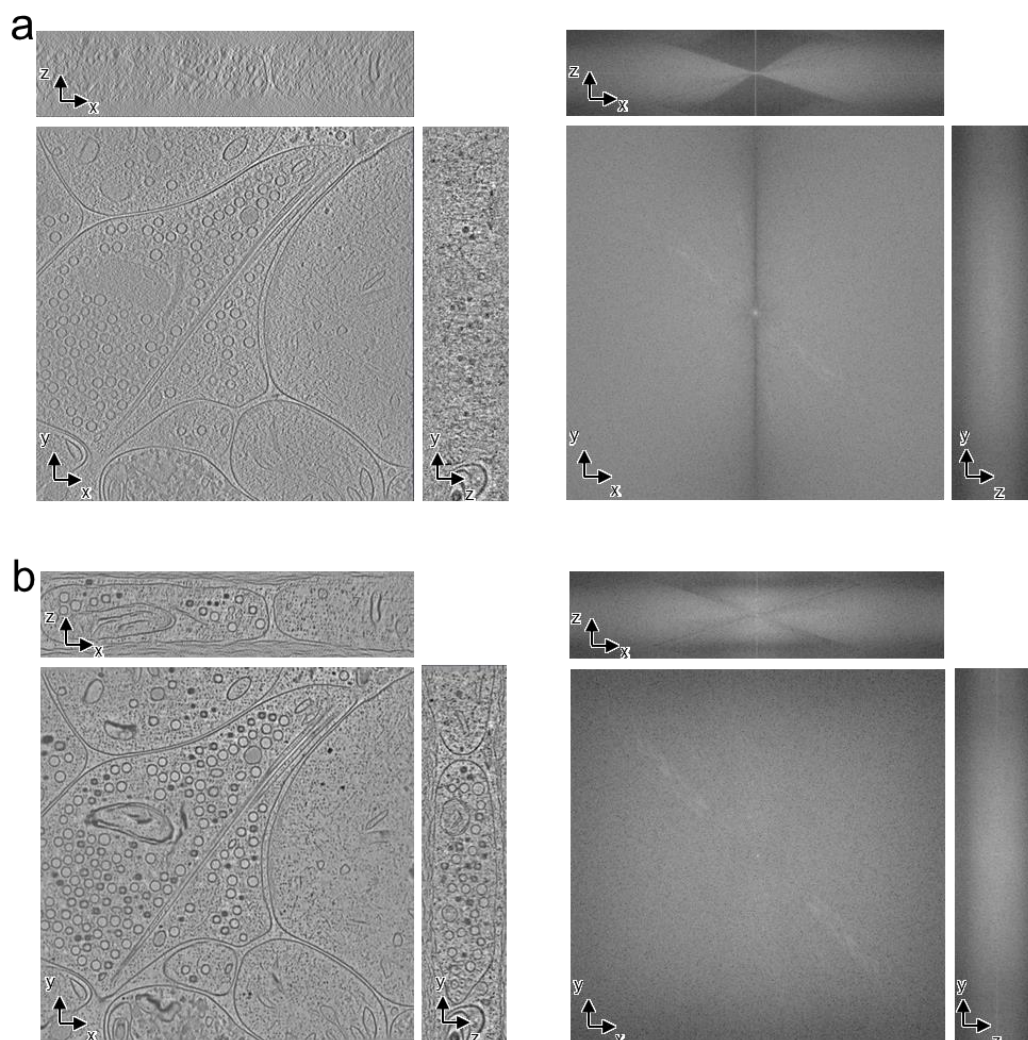
877 **Supplementary Fig. 1│ a,** Orthogonal views of the WBP reconstructed tomograms and their

878 corresponding Fourier transforms. **b,** Orthogonal views of the IsoNet reconstructed

879 tomograms and their corresponding Fourier transforms.

880

881

882
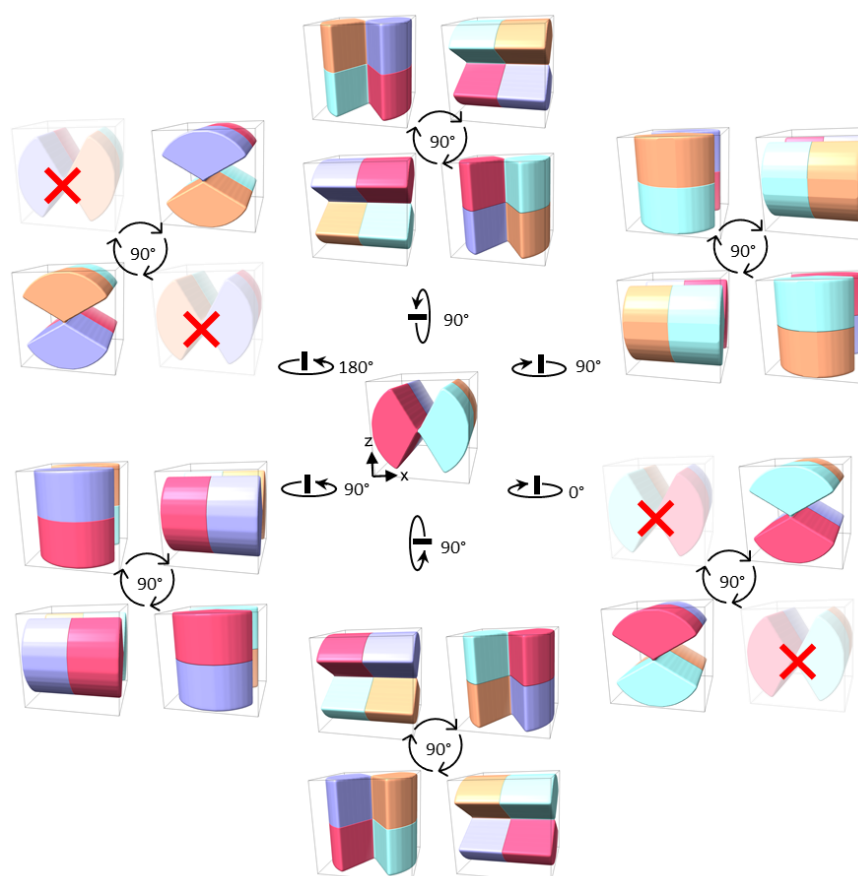
Supplementary Fig. 2 │ Rotation schemes. Twenty rotated copies are obtained for each extracted subtomograms demonstrated in the center. First, each subtomogram has six faces. Each face can be rotated with an out-of-plane angle to face toward six positive directions of the Y-axis. Second, each out-of-plane rotation is followed by four in-plane rotations, making a total of 24 possible rotations. However, among the 24 rotations copies, four have the same Z-axis missing-wedge as the original subtomograms. Thus, these four rotations are excluded (red cross).

890

891

892
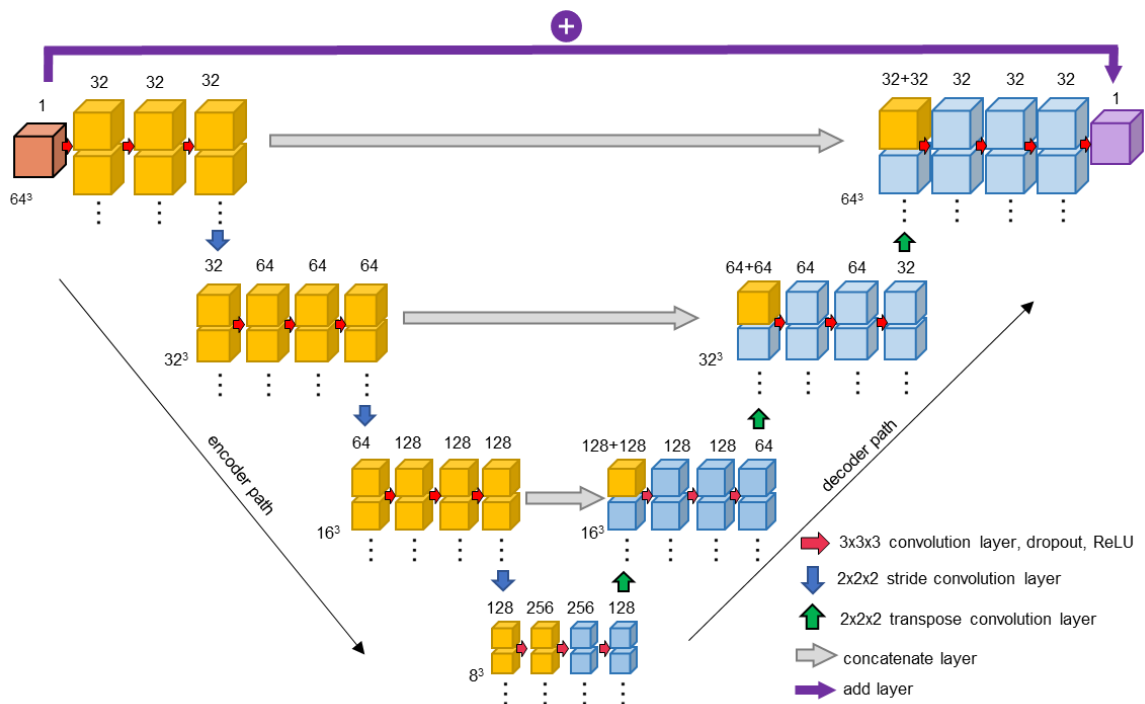
893 **Supplementary Fig. 3 │The architecture of neural network based on U-net.** The values

894 at the bottom left of boxes show sizes of 3D feature maps or subtomograms, while the

895 values on top of the boxes are their numbers.

896

897

898
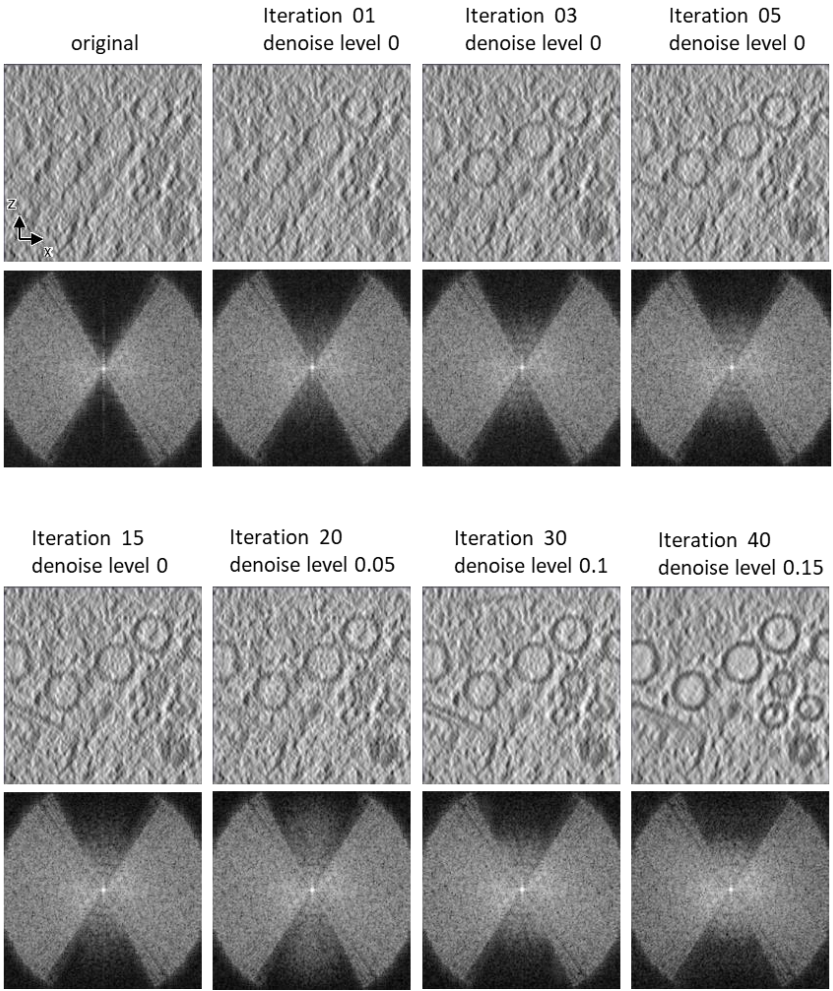
Supplementary Fig. 4 │ Iteratively filling the missing-wedge region. XZ slice views of subtomograms and corresponding power spectrum at different iterations in refine step.

901

902

903

904

905

906
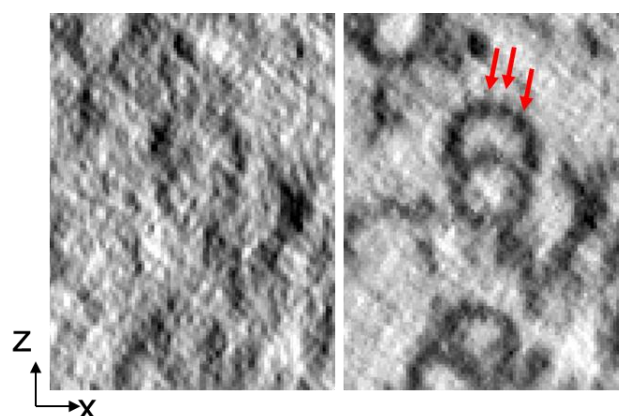
Supplementary Fig. 5 │ XZ slices of microtubule doublets in axoneme. Left: XZ slices of the SIRT reconstructions. Right: the slices of the corresponding tomogram generated by IsoNet. Red arrows indicate microtubule subunits.

910

911
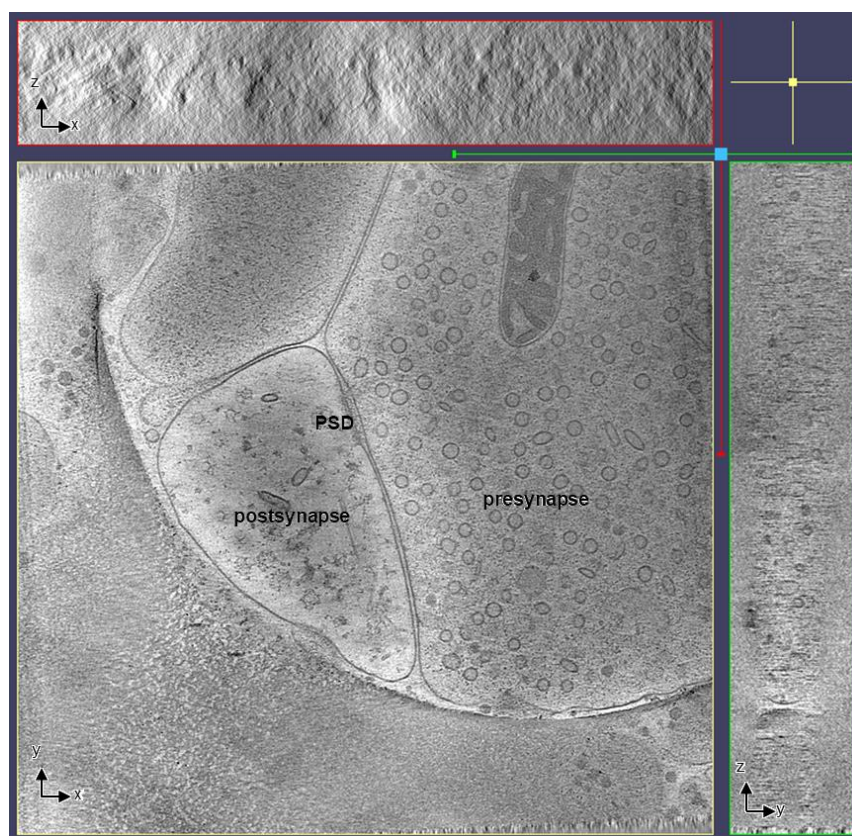
912

913 **Supplementary Fig. 6 | Orthogonal views of the tomogram containing clathrin cages**

914 **reconstructed with the SIRT algorithm in IMOD.** PSD: postsynaptic density.

915

916
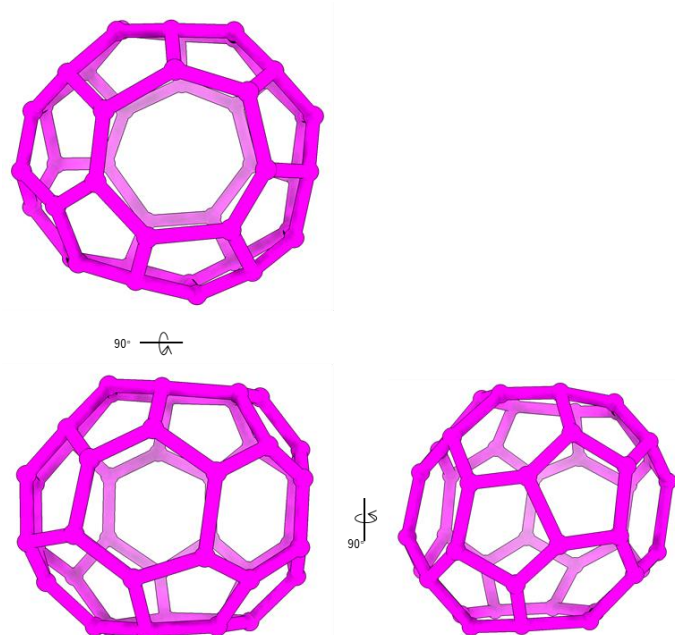
Supplementary Fig. 7 | 3D views for the shape of the heptagon containing clathrin cage

917

918

919
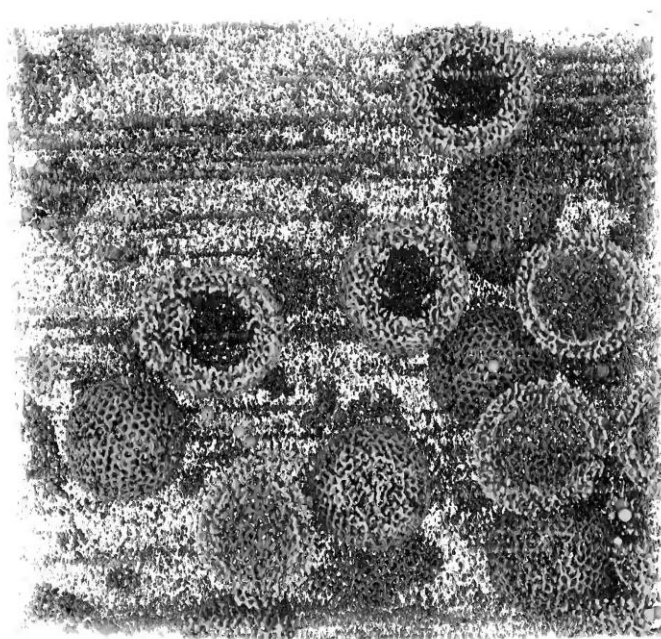
920

921 **Supplementary Video 1 | 3D tomogram of HIV particles.** This video shows 3D structures

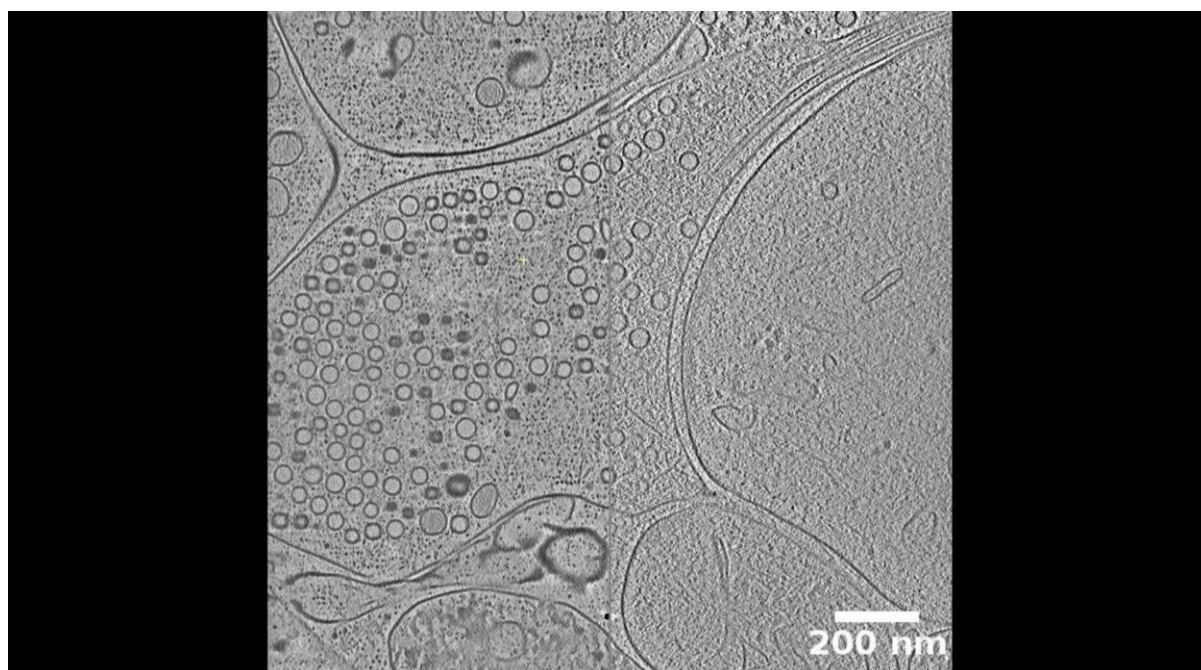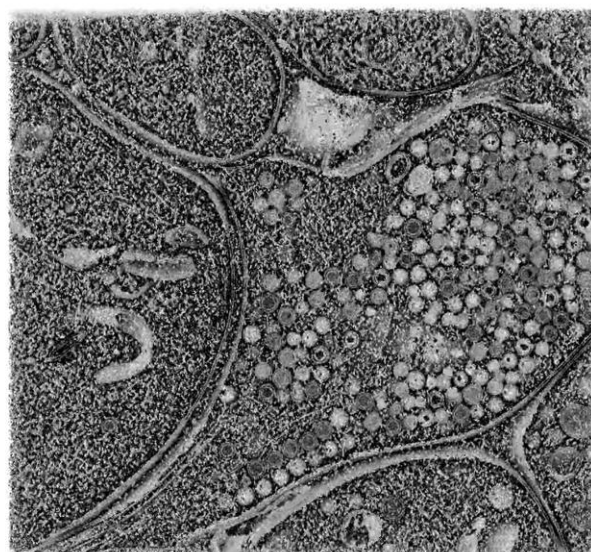922 in an IsoNet generated tomogram. The tomogram density is sliced through three orthogonal

923 directions.

924
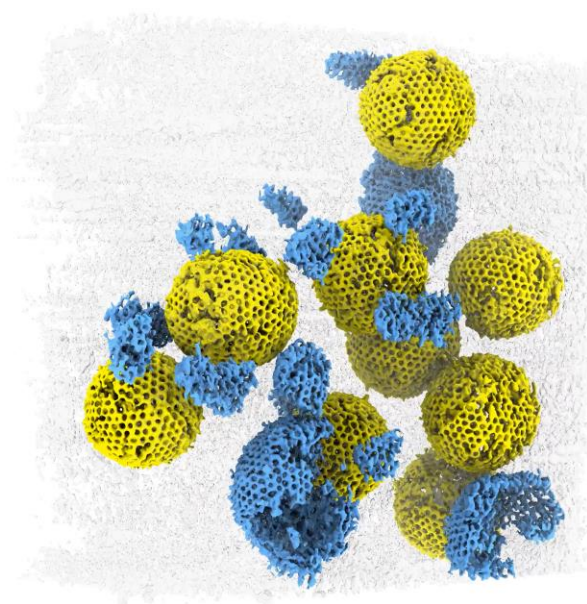
925

926 **Supplementary Video 2 | 3D rendering of HIV particles.** The HIV particles are rendered in

927 yellow (fully embedded in ice) and blue (at air-water interface). The rest cryoEM density is

928 shown in transparent gray.

929

930

Supplementary Video 3 | 2D slices of a tomogram of a neuronal synapse. Left: IsoNet generated tomogram. Right: The original tomogram reconstructed with WBP.
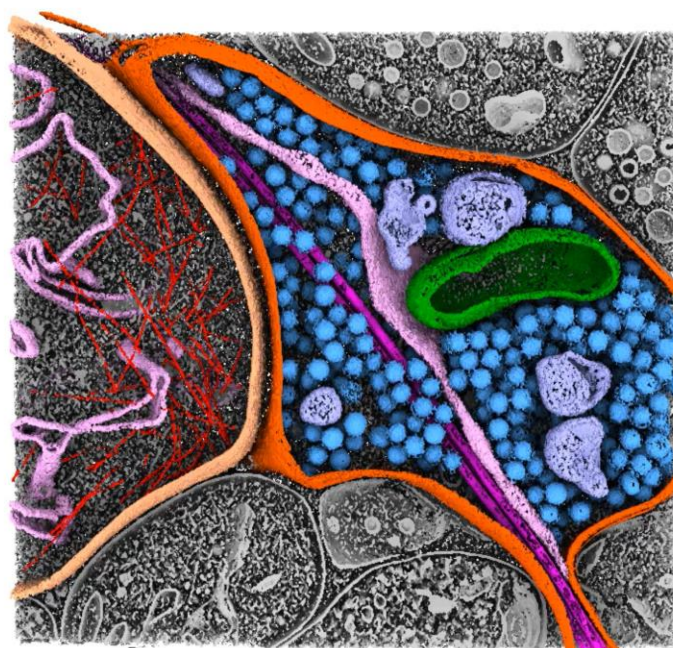
933

934

935 **Supplementary Video 4 | 3D tomogram of a neuronal synapse.** This video shows 3D

936 structures in an IsoNet generated tomogram. The tomogram density is sliced through three
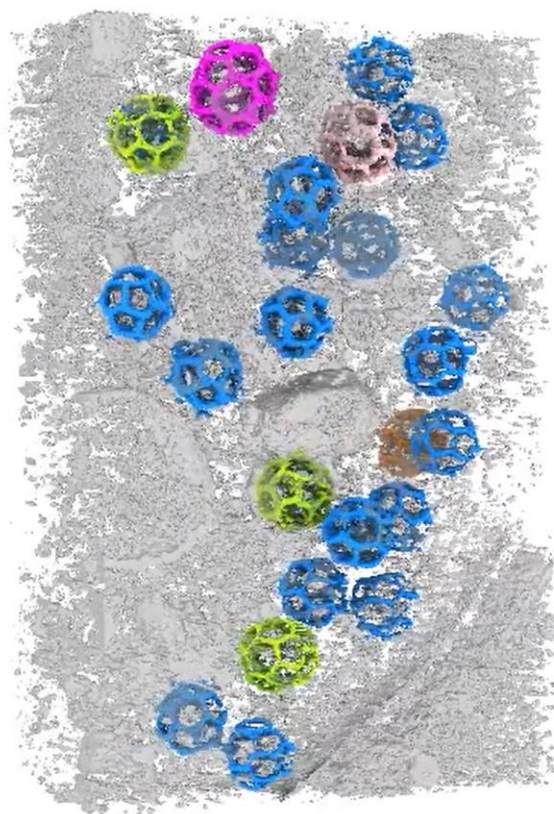
937 orthogonal directions.

938

939

940

Supplementary Video 5 | 3D rendering of the neuronal synapse.

942

943

Supplementary Video 6 | 3D tomogram clathrin cages in a neuronal synapse. This
video shows 3D structures of an IsoNet generated tomogram. The tomogram density is
sliced through three orthogonal directions. Then, 3D rendering of clathrin cages is shown
and rotated.

948

949

950