

# Neural Subdivision

HSUEH-TI DEREK LIU, University of Toronto

VLADIMIR G. KIM, Adobe Research

SIDDHARTHA CHAUDHURI, Adobe Research, IIT Bombay

NOAM AIGERMAN, Adobe Research

ALEC JACOBSON, University of Toronto, Adobe Research

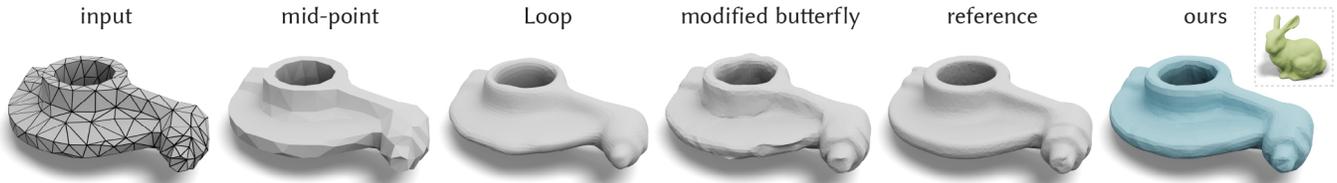


Fig. 1. Our neural subdivision framework performs geometry-aware subdivision, reconstructing the reference rocker arm that we decimated to obtain the coarse input with high accuracy, even though it was only trained on one single model - the Stanford bunny. Neural subdivision does not suffer from the inherent limitations of classic subdivisions, such as volume shrinkage and over-smoothing (Loop [1987]), or amplification of tessellation artifacts (modified butterfly [Zorin et al. 1996]). Throughout this paper, we use green to denote the training shape, and blue for the neural subdivision output.

This paper introduces *Neural Subdivision*, a novel framework for data-driven coarse-to-fine geometry modeling. During inference, our method takes a coarse triangle mesh as input and recursively subdivides it to a finer geometry by applying the fixed topological updates of Loop Subdivision, but predicting vertex positions using a neural network conditioned on the local geometry of a patch. This approach enables us to learn complex non-linear subdivision schemes, beyond simple linear averaging used in classical techniques. One of our key contributions is a novel self-supervised training setup that only requires a set of high-resolution meshes for learning network weights. For any training shape, we stochastically generate diverse low-resolution discretizations of coarse counterparts, while maintaining a bijective mapping that prescribes the exact target position of every new vertex during the subdivision process. This leads to a very efficient and accurate loss function for conditional mesh generation, and enables us to train a method that generalizes across discretizations and favors preserving the manifold structure of the output. During training we optimize for the same set of network weights across all local mesh patches, thus providing an architecture that is not constrained to a specific input mesh, fixed genus, or category. Our network encodes patch geometry in a local frame in a rotation- and translation-invariant manner. Jointly, these design choices enable our method to generalize well, and we demonstrate that even when trained on a single high-resolution mesh our method generates reasonable subdivisions for novel shapes.

Authors' addresses: Hsueh-Ti Derek Liu, University of Toronto, hsuehtil@cs.toronto.edu; Vladimir G. Kim, Adobe Research, vokim@adobe.com; Siddhartha Chaudhuri, Adobe Research, IIT Bombay, sidch@cse.iitb.ac.in; Noam Aigerman, Adobe Research, aigerman@adobe.com; Alec Jacobson, University of Toronto, Adobe Research, jacobson@cs.toronto.edu.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from [permissions@acm.org](mailto:permissions@acm.org).

© 2020 Association for Computing Machinery.

0730-0301/2020/7-ART1 \$15.00

<https://doi.org/10.1145/3386569.3392418>

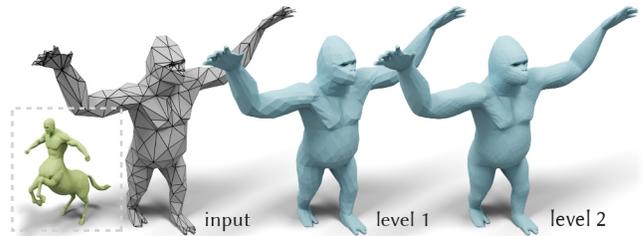


Fig. 2. Neural subdivision refines different parts of a mesh differently, conditioned on the local geometry. Here, the network was trained on the centaur model (green) and then evaluated on a coarse gorilla mesh (gray).

## ACM Reference Format:

Hsueh-Ti Derek Liu, Vladimir G. Kim, Siddhartha Chaudhuri, Noam Aigerman, and Alec Jacobson. 2020. Neural Subdivision. *ACM Trans. Graph.* 39, 4, Article 1 (July 2020), 16 pages. <https://doi.org/10.1145/3386569.3392418>

## 1 INTRODUCTION

Subdivision surfaces are defined by deterministic, recursive up-sampling of a discrete surface mesh. Classic methods work by performing two steps: each input mesh element is *divided* into many elements (e.g., one triangle becomes three) by splitting edges and adding vertices. The positions of the mesh vertices are then smoothed by taking a weighted average of their neighbors' positions according to a weighting scheme based purely on the local mesh connectivity. Subdivision surfaces are well studied and have rich theory connecting their limit surfaces (applying an infinite number of subdivide-and-smooth iterations) to traditional splines. They are a standard paradigm in surface modeling tools, allowing modelers to sculpt shapes in a coarse-to-fine manner. A modeler may start with a very coarse cage, adjust vertex positions, then subdivide once, adjust the finer mesh vertices, and repeat this process until satisfied.

While existing subdivision methods are well-suited for this sort of interactive modeling, they fall short when used to automatically



Fig. 3. Neural subdivision can adapt to different input triangulations and output a high-resolution surface mesh accordingly. This enables us to use it directly in the graphics pipeline such as texture mapping.

upsample a low resolution asset. Without a user’s guidance, classic methods will overly smooth the entire shape (see Fig. 1). Popular methods based on simple linear averaging do not identify details to maintain or accentuate during upsampling. They make no use of the geometric context of a local patch of a surface. Furthermore, classic methods based on fixed one-size-fits-all weighting rules are determined for their general convergence and smoothness properties. This ignores an opportunity to leverage the massive amount of information lurking in the wealth of existing 3D models.

We propose *Neural Subdivision*. We recursively subdivide an input triangle mesh by applying the same fixed topological updates of classic Loop Subdivision, but move vertices according to a neural network conditioned on the local patch geometry. We train the shared weights of this network to learn a geometry-dependent non-linear subdivision that goes beyond classic linear averaging (see Fig. 2). The choice of training data tailors the network to a particular class, type or diversity of geometries.

An immediate challenge is how to collect training data pairs. There is an ever-growing number of 3D models available. However, many if not most of them were not created using a subdivision modeling tool. Even among those that were, the final model does not retain information to replay the modeler’s vertex displacements. In the absence of paired data for a supervised training approach, we propose a novel method to *self-supervise* given only high-resolution surface meshes of arbitrary origin/connectivity at training time. We stochastically generate candidate low-resolution versions of a training exemplar while maintaining a bijective correspondence between their surfaces. This correspondence enables a novel loss function that is more efficient and accurate compared to existing methods. By construction, this training regime ensures *generalization across discretization*.

In contrast to existing generative models for surfaces, our output is a surface mesh with deterministic connectivity based on the input, enabling direct use in the standard graphics pipeline such as texture mapping (see Fig. 3). By sharing weights and training across all local patches of all the training meshes, we learn a rule based on the local neighborhood rather than the entire shape. Compared to existing methods, this frees our network from being constrained to a fixed genus, relying on a template, or requiring an extremely large collection of shapes during training. We demonstrate that even when trained on a single shape, our method can generalize to novel meshes. We design our network to encode vertex position data in a local frame ensuring rotation and translation invariances without resorting to handcrafted predefined feature descriptors.

We demonstrate the effectiveness of our method with a variety of qualitative and quantitative experiments. Our method generates subdivided meshes that are closer to the true high-resolution shapes than traditional interpolatory and non-interpolatory subdivision methods, even when trained with a small number of very dissimilar exemplars. We introduce a quantitative benchmark and show significant gains over classic subdivision methods when measuring upsampling fidelity. Finally, we show prototypical applications of Neural Subdivision to low-poly mesh upsampling and 3D modeling.

## 2 RELATED WORK

Our work builds directly upon the foundations of classic subdivision surfaces and connects to the rapidly advancing field of neural geometry learning. We focus this section on establishing context with past subdivision schemes and contrasting our geometric learning contributions with contemporary works.

### 2.1 Subdivision Surfaces

The basic idea of subdivision is to “define a smooth curve or surface as the limit of sequence of successive refinements” [Zorin et al. 2000]. This broad definition admits a wide variety or “zoo” of different subdivision schemes that would be outside the scope of this paper to cover thoroughly. The history of subdivision surfaces reaches back to the early work on irregular polygon meshes [Doo 1978; Doo and Sabin 1998] and the now ubiquitous Catmull-Clark subdivision which produces quad meshes [Catmull and Clark 1998]. The linear method of Loop [1987] for triangle meshes has reached similar popularity, and is the basis for our non-linear neural subdivision.

Classic linear subdivision methods are defined by a combinatorial update (splitting faces, adding vertices, and/or flipping edges [Kobbelt 2000]) and a vertex smoothing (repositioning step) based on local averaging of neighboring vertex positions. Subdivision methods are well studied from a theoretical perspective on the existence, direct evaluation, and continuity of the limit surface [Karciauskas and Peters 2018; Stam 1998; Zorin 2007]. Modelers typically manipulate a subdivision surface in a coarse to fine fashion. Most modeling tools already visualize the limit surface or some approximation of it, while the user manipulates the coarse level (cage) (see Fig. 23). Beyond moving vertices, users can control the surface by adding creases (sharp edges) [DeRose et al. 1998; Hoppe et al. 1994]. Non-interpolating methods such as Catmull-Clark or Loop appear to be the most popular, but interpolating methods do exist (e.g., [Dyn et al. 1990; Kobbelt 1996; Zorin et al. 1996]) and have similar smoothness guarantees, although *fairness* is harder to achieve (see Fig. 1). Linear methods are easier to analyze and design to guarantee smoothness. As a result, capturing details is left to the modeler or a deterministic procedural routine (e.g., [Tobler et al. 2002a,b; Velho et al. 2002]).

Our neural subdivision acts similar to non-linear subdivision methods, with the subdivision rule in this case being a non-linear function learned by a neural network. Non-linear subdivision has been studied from the mathematical perspective [Floater and Michelli 1997; Schaefer et al. 2008] and also as a mechanism to maintain certain geometric invariants during each level of subdivision (e.g., circle-preserving [Sabin and Dodgson 2004], quad planarity [Bobenko et al. 2020; Liu et al. 2006], developability [Rabinovich

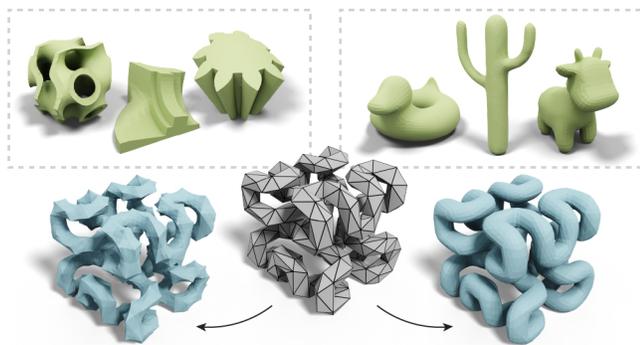


Fig. 4. Our subdivision is data-driven. Training on a set of mechanical objects (left, green) or a set of smooth organic objects (right, green) leads to drastically different styles (blue). ©Gyroid Puzzle by eemmett (top left) and Hilbert Cube by tbuser (bottom) under CC BY-SA.

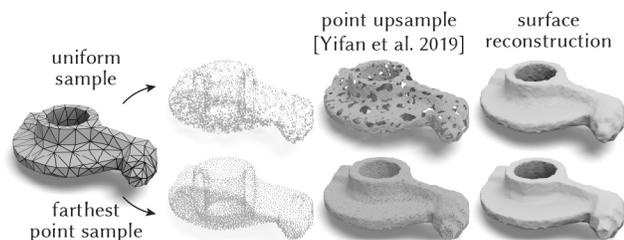


Fig. 5. One can use existing point upsampling methods to refine coarse meshes by (1) sampling, (2) upsampling [Wang et al. 2019b], and (3) reconstruction [Kazhdan and Hoppe 2013]. However, this may lead to artifacts since it lacks information about the surface, and requires the use of expensive surface reconstruction as a post-process.

et al. 2018; Tang et al. 2014], Möbius-regularity [Vaxman et al. 2018], cloth wrinkliness [Kavan et al. 2011]). One general approach is to combine a linear subdivision with an online geometric optimization, and recursively apply the non-linear rule an arbitrary, if not infinite number of times, akin to classic linear rules. Our approach can be viewed as an extreme form of precomputation, where the optimization is the training procedure and the fixed network is applied generally as a non-linear function evaluation. The choice of data in the training will influence the “style” of our non-linear subdivision (see Fig. 4). Although our method is non-linear, it is trained to work well for a pre-specified finite number of times.

Recently, Preiner et al. [2019] introduced a new non-linear subdivision method that treats the coarse shape probabilistically. Their contributions are orthogonal to ours, and while we base our method on Loop subdivision, we could in theory extend our network to learn on top of this more powerful subdivision method.

## 2.2 Neural Geometry Learning

Recent advances in generative neural networks enabled the use of learnable components in 3D modeling applications such as shape completion [Li et al. 2019], single-view [Tatarchenko et al. 2019] and multi-view [Sitzmann et al. 2019] reconstruction, and modeling-by-parts [Chaudhuri et al. 2020].



Fig. 6. We compare the same model trained using (a) chamfer distance (which only measures error between point sets) and (b) our  $\ell^2$  loss based on shape correspondences. The model trained using the chamfer distance fails to capture the surface topology (red). In contrast, our loss function leads to manifold output meshes (blue). ©Gyroid Puzzle by eemmett (top) under CC BY-SA.

The closest to our neural mesh subdivision application are the deep point cloud upsampling techniques [Li et al. 2019; Wang et al. 2019b; Yu et al. 2018]. The disadvantage of using a point cloud as input is that it lacks connectivity information, and requires the neural network to implicitly estimate the structure of the underlying manifold. Meshes can also be more efficient at representing feature-less regions with larger planar elements, providing a wider reception field to our mesh-based neural network. Mesh output is preferred for many standard graphics pipelines, thus, a post process is often required [Kazhdan and Hoppe 2013] to convert the output of point-based methods to meshes, which prevents building an end-to-end trainable system. Fig. 5 illustrates the output of a point upsampling method that was pre-trained on a collection of statues [Wang et al. 2019b] (see App. A for implementation details).

Our work is related to other neural mesh generation techniques. Free-form generation of meshes as a set of vertices and faces is infeasible with current deep learning methods, due to the lack of regular structure, uneven discretization, and combinatorial variability in the possible outputs, limiting such approaches to very coarse outputs [Dai and Nießner 2019]. A common alternative is to deform a global template either by predicting vertex coordinates [Ranjan et al. 2018; Tan et al. 2018] or by training a deformation network that warps the entire 3D domain conditioned on a latent vector that encodes the deformation target [Groueix et al. 2018a; Yifan et al. 2020]. While these approaches usually produce meshes with higher resolution, their output is limited to deformations of a single shape. Some techniques propose using generic templates such as spheres [Wang et al. 2018; Wen et al. 2019] or 2D atlases [Groueix et al. 2018b], which place limitations on the topology of the output. In contrast to these techniques, our method refines the mesh locally, and thus, respects the topology of the input (which could be arbitrary). Another advantage of our local refinement approach is that we do not require co-aligned training data with a well-defined object space, the output of our subdivision networks is translation and rotation invariant since it can be described in a local coordinate system of the input patch.

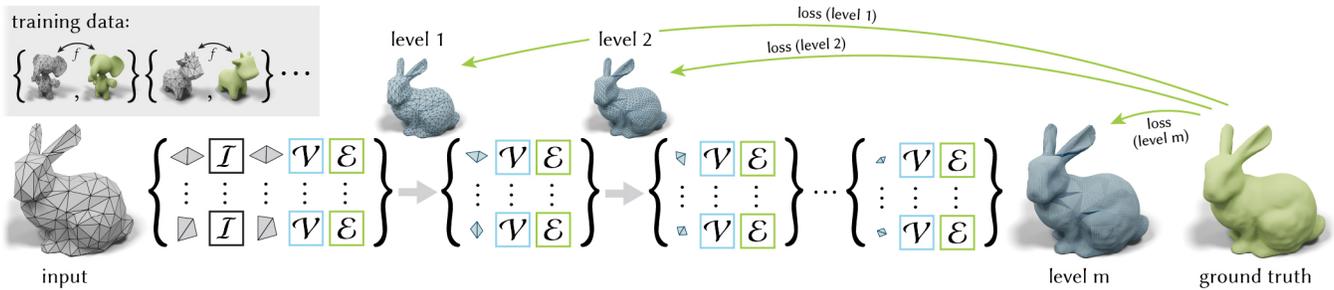


Fig. 7. Neural subdivision takes a coarse triangle mesh (gray) as input and outputs a sequence of subdivided meshes (blue) with different levels of details. During training, we minimize the  $\ell^2$  loss from the ground truth (green) to the output meshes (blue) across levels. Our training data consists of pairs of coarse and fine meshes (top left) with a bijective map  $f$  between each pair.

There are several options for analyzing a mesh patch with a neural network, such as using a local [Masci et al. 2015; Poulenard and Ovsjanikov 2018] or global [Maron et al. 2017] parameterization to unfold a mesh into 2D grid, or apply graph-based techniques adapted for meshes [Kostrikov et al. 2018; Wang et al. 2019a] (see [Bronstein et al. 2017] for a more comprehensive survey). Our approach is inspired by MeshCNN [Hanocka et al. 2019]. Their method directly learns filters over the local mesh structure via undirected edges, and shows applications in deterministic tasks. In contrast, we focus on generative tasks and develop a novel set of features over the *half-flaps* – an edge along with its two adjacent triangles. Each half-flap has a canonical orientation which gives us well-defined local frames which are crucial for our network’s rigid motion invariance.

Geometry generation techniques are typically trained with reconstruction losses that measure how well does the generated surface approximate the known target. Surface-to-surface distances are commonly employed, with correspondences defined via closest-point queries (aka chamfer distance) [Barrow et al. 1977; Fan et al. 2017]. However, the closest-point approach matches many points to the same point, while leaving other points unmatched, resulting in self-overlaps and unrepresented areas (see Fig. 6).

Indeed, prior work demonstrates that using higher quality correspondence (e.g., ground truth mapping) significantly improves results [Groueix et al. 2018a]. While the latter is not available in our setting, we propose a **data generation technique** for creating various coarse variants of the same high-res mesh with a low-distortion bijective map. Bijectivity is crucial for the quality of our training data, ensuring no self-overlaps exist and that every part of the target surface has a pre-image on the coarse mesh.

### 3 NEURAL SUBDIVISION

In the following we overview the main components of our neural subdivision: the test-time inference pipeline, training and loss (Sec. 4), data-generation (Sec. 4), and finally the network architecture (Sec. 5). Later sections will discuss these components in detail.

*Inference.* As illustrated in Fig. 7, our method takes a coarse triangle mesh (gray) as input and recursively refines it by subdividing each triangle to create additional vertices and faces. The output is a sequence of subdivided meshes (blue) with different levels of details. Our subdivision process follows a simple topological update rule (same as Loop), namely inserting new vertices at the midpoints of

all edges. It then uses a neural network to predict new positions for all vertices, at each new level of subdivision.

*Training and loss.* The data we generate provides us with correspondences between predicted vertices and points inside the triangle on the ground truth shape. We train our network with the simple  $\ell^2$  loss, by measuring the distance between each predicted vertex position at every level of subdivision and its corresponding point on the original shape (green). As there is no existing dataset consisting of pairs of high-quality meshes and subdivision surfaces in correspondence, we instead develop a novel technique for generating training data, comprising of coarse and fine meshes with bijective mappings between them.

*Data generation.* We first note that each vertex  $v$  created from a subdivision step has a well-defined mapping back to the coarse mesh, defined by mapping that vertex to its corresponding midpoint. Thus, each subdivided mesh at any level of subdivision can be mapped back to the initial coarse mesh via a sequence of mid-point-to-vertex or vertex-to-vertex maps. In practice we use barycentric coordinates to encode this **subdivided-to-coarse bijective mapping**,  $g$ . Hence, if we had a bijective mapping  $f$  between the coarse mesh and the original mesh, we could define a unique point on the original mesh corresponding to  $v$ , by compositing the two maps:  $f(g(v))$ .

Thus, the only missing part is to create coarse and fine meshes with bijective mappings between them. We achieve this by taking a high-resolution training mesh and sequentially coarsening it by applying random sequences of edge collapses, thereby generating a sequence of coarsened meshes. We maintain low-distortion correspondences between the coarsened and original mesh by computing a conformal map between the 1-ring edge neighborhood (before the collapse) and the 1-ring vertex neighborhood (after the collapse). Composition of these maps creates a dense bijective map  $f$  between the coarse and original meshes, which is then directly applied to the training (Fig. 9).

*Advantages of our training approach.* In comparison to closest point losses that are commonly used to train generative neural networks, our correspondence-based loss is aware of the manifold structure (Fig. 6) and is orders of magnitude faster to compute (Fig. 17). Bijectivity and continuity of the map ensure that the entire ground truth surface is captured by some region of the coarse mesh (Fig. 10). The low distortion encourages uniformity, which in

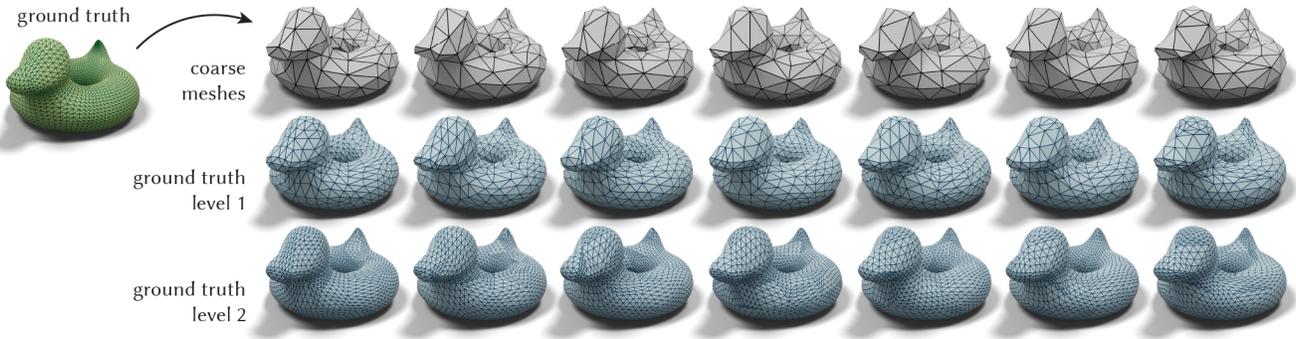


Fig. 8. Given a ground truth shape (green), we use random edge collapses to create several coarse meshes (gray). For each coarse mesh, we subdivide the mesh and use the bijective map to determine the position on the ground truth for all the vertices across different levels. The blue meshes are the ground truth subdivisions that exhibit one-to-one vertex correspondences to the network predictions.



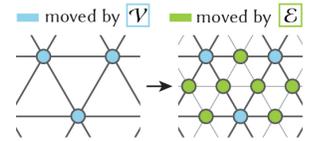
Fig. 9. Given an edge collapse algorithm of choice, we plug in our successive self-parameterization described in Sec. 4.1 to compute a bijective map between the original mesh (green) and its decimated version (gray). We visualize the map by coloring the fine mesh using the triangulation of the coarse mesh (right). ©Tarbosaurus Skull by gpvillamil under CC BY-SA.

turn enables the reproduction of the target surface with just a few uniform subdivisions, and, more importantly reduces the variance in the signal the network needs to learn. We can further leverage the low-distortion map to map an additional signal, such as texture (Fig. 3). As our training data contains many pairs with different random decimations of the same ground truth (App. E), our network is able to learn how to generalize across discretization.

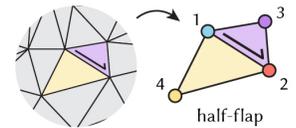
**Network architecture.** Similarly to the subdivision process, the learnable modules of our network are applied recursively. They operate over atomic local mesh neighborhoods and predict differential features (meaning they represent geometry in the local coordinates of the mesh, and not in world coordinates). These features are then used to compute vertex coordinates at the new level of subdivision. We define three types of modules applied at three sequential steps. During the **INITIALIZATION step**, we first compute differential per-vertex quantities that are based on the local coordinate frame. A learnable module  $\mathcal{I}$  is applied to the 1-ring neighborhood of every vertex to map these differential quantities to a high-dimensional feature vector stored at the vertex. Note that this high-dimensional feature vector is a concatenation of a learnable latent space which *encodes* local geometry of the patch, and differential quantities which

directly *represent* local geometry and enable us to reconstruct the vertex coordinates.

For each subsequent subdivision iteration, we assume that the topology is updated following the Loop subdivision scheme, splitting each edge at midpoint, and consequently, subdividing each triangle into four (see inset). A **VERTEX step** uses the module  $\mathcal{V}$  to predict vertex features for the next level of subdivision based on its 1-ring neighborhood, where vertices affected by this step only involve corners of triangles at the previous mesh level. Then, an **EDGE step** uses the module  $\mathcal{E}$  to compute features of vertices added at midpoints based on the pair of vertices that were connected by an edge at the previous mesh level.



Our modules share a very similar architecture and heavily rely on a learnable operator defined over a *half-flap*: a directed edge and its two adjacent triangles (see the inset). We use the directed edge to define the local coordinate frame which is used to estimate the differential features of either input or output of learnable modules. Note also that the directed edge allows us to order the four adjacent vertices of the flap in a canonical way. We concatenate their features and feed them into **shallow multi-layer perceptrons** (MLP). The weights of the MLPs are shared within each module type and across all levels of subdivision. Both modules  $\mathcal{I}$  and  $\mathcal{V}$  process all half-flaps defined by an outgoing edge and use average pooling to combine the half-flap features into per-vertex features. The module  $\mathcal{E}$  also combines features from two half-flaps (both directions of the edge) via average pooling. Since our architecture is local, and uses input and output features that are invariant to rigid motions, it exhibits an impressive ability to generalize from example, even when trained on a single fine mesh.



## 4 DATA GENERATION AND TRAINING

While our network architecture and invariant layers are crucial for its ability to learn subdivisions, it by its own is only half of the two main components that together facilitate high-quality neural

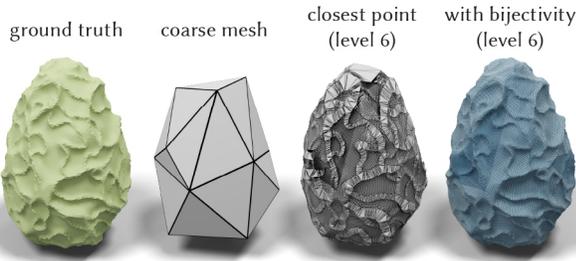


Fig. 10. Given a ground truth/coarse mesh pair, naively using “closest-point-on-mesh” to estimate correspondences between the level-6 subdivided mesh and the ground truth results in a non-bijective map, causing the loss function to fail to capture the entire ground truth mesh (third column). Our successive self-parameterization ensures bijectivity, which implies the entire ground truth surface will be captured (right)

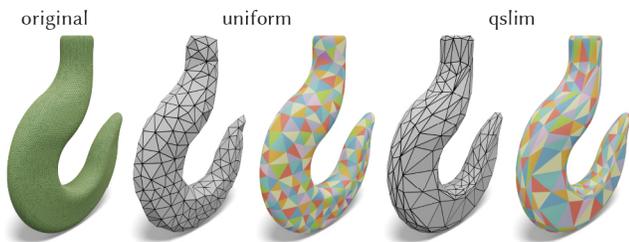


Fig. 11. Different edge-collapse algorithms can be used in a plug-and-play manner to create, for instance, a uniform-area parameterization (middle) and an appearance-preserving parameterization (right). This flexibility is used to create training data with diverse types of discretizations.

subdivisions. The other half consists of the training process, data and the loss function.

Consider a naive approach to the subdivision training: generate pairs of coarse/fine meshes by a decimation algorithm; measure the distance between the network’s predicted subdivision and the ground truth, for instance by the average distance between predicted points and their projections on the ground truth mesh; iterate over coarse/fine pairs while optimizing the loss. This naive approach has a major caveat. Computing correspondences using the chamfer-like loss (Fig. 6) or point-to-mesh distance (Fig. 10) is known to lead to incorrect and self-overlapping matches between shapes. This leads to a badly training set up because the loss itself exhibits artifacts.

In lieu of this naive approach, we consider the fact that a pair of coarse and fine meshes both approximate the same underlying smooth surface. This motivates us to compute the correspondences based on the intrinsic geometry, instead of an ad-hoc correspondence. The outcome is a high-quality bijective map between each pair of coarse and fine meshes, enabling us to obtain one-to-one vertex correspondences. Therefore a simple  $\ell^2$  loss is sufficient to correctly measure the error between every level of neural subdivision and the ground truth shape.

#### 4.1 Successive Self-Parameterization

One possible solution is to apply general shape matching techniques. But ensuring bijectivity in general shape matching is difficult. For

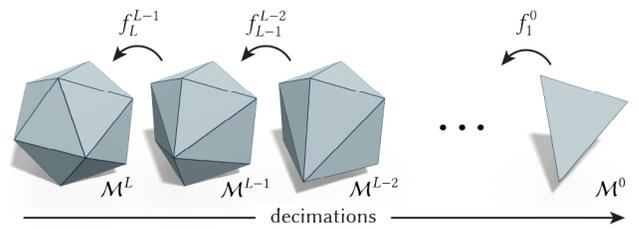


Fig. 12. We compute a bijective map for each edge collapse. The bijective map from the coarsest mesh  $M^0$  to the input mesh  $M^L$  is then computed by composing all the maps  $f_l^{l-1}$ .

instance, it requires the two shapes to have the same number of vertices [Vestner et al. 2017], or a user-guided common domain [Praun et al. 2001; Schreiner et al. 2004], or user-provided landmark correspondences [Aigerman et al. 2014, 2015; Kraevoy and Sheffer 2004] (see [Van Kaick et al. 2011] for a survey). However, our problem is considerably simpler, since we aim to construct a map between different discretizations of the same shape, and we have full control on the decimation procedure.

The closest solution to our problem is a seminal work – **MAPS** [Lee et al. 1998] – on self-parameterization. Given an input mesh, MAPS computes the bijective map by successively removing vertices of the maximum independent set. Since then, several improvements have been proposed [Guskov et al. 2002, 2000; Khodakovskiy et al. 2003]. Unfortunately, they cannot be directly applied to edge collapses for creating training data for our learning task (see App. D). We need an algorithm that has the flexibility to be used with any edge decimation method, so that we can generate a diverse collection of coarse meshes (see Fig. 11). Fortunately, the idea from [Cohen et al. 1998, 1997, 2003] for minimizing mesh/texture deviation leads us to generalize the idea of MAPS to any edge collapses.

Our method for computing the bijective map, designed specifically for creating data to train neural subdivision, combines the idea of self-parameterization from MAPS [Lee et al. 1998] and the idea of successive mapping from [Cohen et al. 1998, 1997, 2003]. Thus, we call it **successive self-parameterization**. This combination enables us to compute the parameterization intrinsically to avoid the requirement of having a given UV map, such as in the method of Liu et al. [2017]. The result of the combination is extremely simple. It is a two-step module that can be applied to any choice of edge-collapse algorithm (see Fig. 11) and it will output a bijective map after the decimation. Hence, the inputs to successive self-parameterization are a triangle mesh and an edge collapse algorithm of choice, and the output is a decimated mesh with a corresponding bijective map between the input and the decimated model. For the sake of reproducibility, we reiterate the core ideas from [Cohen et al. 1998, 1997; Lee et al. 1998], and describe how to combine both ideas.

We denote the input triangle mesh as  $M^L = (V^L, F^L)$ , where  $V^L, F^L$  are vertex positions and face information respectively at the original level  $L$ . The input mesh  $M^L$  is successively simplified into a series of meshes  $M^l = (V^l, F^l)$  with  $0 \leq l \leq L$ , where  $M^0 = (V^0, F^0)$  is the coarsest mesh. For each edge collapse  $M^l \rightarrow M^{l-1}$ , we compute the bijective map  $f_l^{l-1} : M^{l-1} \rightarrow M^l$  (see Fig. 12) on the

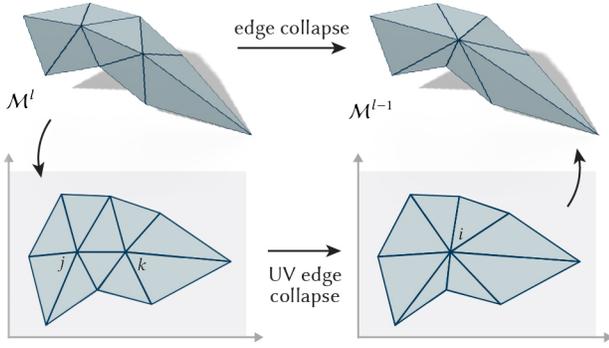


Fig. 13. For each edge collapse, we simultaneously collapse the edge on the 3D mesh (top) and the UV domain (bottom). As the boundary vertices of the edge’s 1-ring are preserved through the edge collapse, we constrain the flattened boundary in UV space to be at the same position when computing the conformal parameterization of the post-collapse 1-ring.

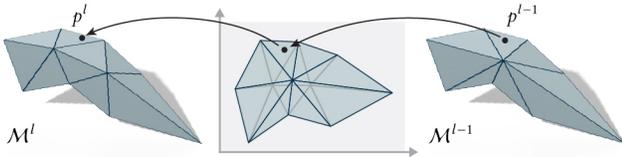


Fig. 14. Since both the pre-collapse and post-collapse parameterizations of the 1-ring map it into the same 2D domain, we can easily use the shared UV space to map a point back and forth between  $M^l$  and  $M^{l-1}$ .

fly. The final map  $f_L^0 : M^0 \rightarrow M^L$  is computed via composition,

$$f_L^0 = f_L^{L-1} \circ \dots \circ f_1^0. \quad (1)$$

We now focus our discussion on the computation of a bijective map for a single edge collapse.

## 4.2 Single Edge Collapse

In each edge collapse, the triangulation remains the same, except for the neighborhood of the collapsed edge. Let  $\mathcal{N}(i)$  be the neighboring vertices of a vertex  $i$  and let  $\mathcal{N}(j, k) = \mathcal{N}(j) \cup \mathcal{N}(k)$  denote the neighboring vertices of an edge  $(j, k)$ . After each collapse, the algorithm computes the bijective map for the edge’s 1-ring  $\mathcal{N}(j, k)$ , in two stages. It first parameterizes the neighborhood  $\mathcal{N}(j, k)$  (prior to the collapse) into 2D. It then performs the edge collapse both on the 3D mesh, and in UV space, as depicted in Fig. 13. The key observation from [Cohen et al. 1998, 1997] is that the boundary vertices of  $\mathcal{N}(j, k)$  before the collapse become the boundary vertices of  $\mathcal{N}(i)$  after the collapse. Hence the UV parameterization of the 1-ring remains valid and injective after the collapse. Then, for any given point  $p^{l-1} \in M^{l-1}$  (represented in barycentric coordinates), we can utilize the shared UV parameterization to map  $p^{l-1}$  to its corresponding barycentric point  $p^l \in M^l$  and vice-versa, as shown in Fig. 14.

Following the idea of **MAPS** [Lee et al. 1998], we use conformal flattening [Mullen et al. 2008] to compute the UV parameterization of the 1-rings, Fig. 13. After collapsing an edge and inserting the new vertex  $v \in \mathbb{R}^3$ , we determine this vertex’s UV location by performing

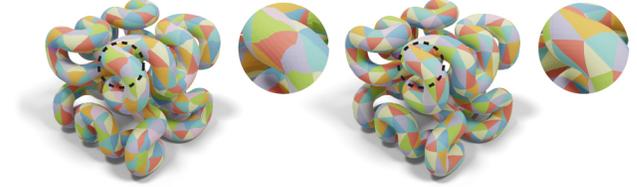


Fig. 15. Using a different parameterization technique that does not result in a conformal flattening leads to a distorted parameterization (left), in contrast to the conformal parameterization we use, that reduces the amount of angle distortion accumulated throughout the edge collapse sequence (right). ©Hilbert Cube by tbusser under CC BY-SA.

another conformal flattening with fixed boundary. The conformality of the map is crucial, as it minimizes angle distortion which would otherwise accumulate throughout the successive parameterizations, leading to distorted, skewed correspondences and hindered learning of the network (see Fig. 15).

## 4.3 Implementation

Successive self-parameterization can be used with any edge collapse algorithm simply by adding two additional steps (see App. B). The actual edge collapse algorithm, such as QSLIM [Garland and Heckbert 1997], takes  $O(N \log N)$  time, and the flattening is a constant cost on top of each collapse (assuming valence is bounded). Thus the complexity of the entire algorithm containing both edge collapses and successive self-parameterization is still  $O(N \log N)$ .

The robustness of the parameterization algorithm relies heavily on the robustness of the underlying edge collapse algorithm. Edge collapses that may lead to self-intersections can result in unusable maps. In App. C, we summarize our criteria for checking the validity of an edge collapse. This is crucial to ensure that we can generate training data using a wide range of shapes (see Fig. 16).

## 4.4 Training Data & Loss Computation

Our training data is constructed by applying the successive self-parameterization on top of random edge collapses. In Fig. 8, given a high-resolution shape (green), we use QSLIM [Garland and Heckbert 1997] with a random sequence of edge collapses to construct several different decimated models (gray). During the collapse, we plug in our self-parameterization to obtain a high-quality bijective map for each coarse and fine pair.

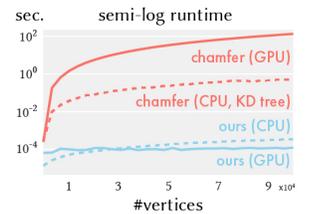


Fig. 17. Our loss computation is orders of magnitude faster than the chamfer loss on the GPU (KAOLIN [J. et al. 2019]) or the CPU (our KD-tree-based implementation).

After the network subdivides the coarse mesh, we use the map to retrieve one-to-one correspondences to the input shape. Specifically, when retrieving the correspondences, we use the Loop topology update to add points in the middle of each edge, e.g., the point with barycentric coordinates  $(0.5, 0.5, 0)$  in a triangle of the coarse mesh. We use these barycentric coordinates  $b$  on the coarse mesh to obtain



Fig. 16. Checking the criteria of collapsible edges is crucial for the robustness of the successive self-parameterization. From left to right, ©Psycho by Aeva (2nd, CC BY-SA), Parametric Sculpture by MCompeau (4th, CC BY-NC), Deer Head by TakeshiMurata (5th, CC BY-SA), Brain Slug by Zarquon (6th, CC BY-NC-SA), Spiral Light Bulb by benglish (7th, CC BY-SA), and Metatron by addy (9th, GNU).

the barycentric coordinates  $f(b)$  on the fine mesh, as illustrated in Fig. 14 using the bijective map  $f$ . During training, suppose  $\mathcal{E}(b)$  is the vertex position output by the network  $\mathcal{E}$ . We measure the per-vertex loss with the  $\ell^2$  distance  $\|f(b) - \mathcal{E}(b)\|_2$ . Compared to the chamfer distance [Barrow et al. 1977], a widely used distance in training 3D generative models [Fan et al. 2017], our loss computation is orders of magnitude faster (see Fig. 17).

## 5 NETWORK ARCHITECTURE

Given a mesh at a previous level of subdivision along with a known topological update rule (mid-point subdivision as used by Loop), our neural network computes all vertex coordinates for the subdivided mesh. Our process involves three main steps illustrated in Fig. 18. The **INITIALIZATION** step uses a learnable neural module  $\mathcal{I}$  to map input per-vertex features to high-dimensional feature vector at each vertex. In each subdivision iteration, the **VERTEX** step uses a learnable module  $\mathcal{V}$  to update features at corners of triangles of the input mesh, and the **EDGE** step uses a learnable module  $\mathcal{E}$  to compute features of vertices that were generated at mid-points of edges of the input mesh. Our network is inspired by classical subdivision algorithms which have two sets of rules: to update (1) *even* vertices from previous iterations, and (2) the newly inserted *odd* vertices. One difference of our approach is that we apply  $\mathcal{V}$  and  $\mathcal{E}$  in sequence, instead of in parallel. This allows us to harness neighborhood information from previous steps.

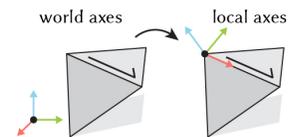
We make several design choices that are critical to the ability of our network to generalize well even from very small amount of training data. First, even though all mesh update steps are global (i.e., they affect every vertex of the mesh), our learnable modules that are used in these steps operate over local mesh patches and share weights. Thus, even a single training pair provides many local mesh patches to train our neural modules. Second, our modules operate over original discrete elements of the mesh, and do not require re-parameterizing or re-sampling the surface. Representing input and output using the mesh discretization enables us to preserve the topology of the input, and generalize to novel meshes with different topology. Third, we represent our vertices using differential quantities with respect to a local coordinate frame instead of using global coordinates. Thus our neural modules operate over

representation that is invariant to rigid motion which simplifies training and improves their ability to generalize.

The key component of our neural module is a **learnable operator** that takes **half-flap**, a 2-face flap adjacent to a half-edge, inspired by the edge convolution approach of Hanocka et al. [2019]. We choose to use half-flap (instead of a flap around an undirected edge) since it provides a unique canonical orientation for the four vertices at the corners of adjacent faces. It also provides a well-defined local coordinate frame which we will use to define differential vertex quantities for the input and output (see the inset). Each flap operator is a shallow multi-layer perceptron (MLP) defined over features of four ordered points. We train one operator per module ( $\mathcal{I}$ ,  $\mathcal{V}$ ,  $\mathcal{E}$ ) across all levels of subdivision and training examples.

Equipped with the half-flap operator, we use average pooling to aggregate features from different half-flaps to per-vertex features in all our neural subdivision steps. **INITIALIZATION** and **VERTEX** steps apply the half-flap operator to every outgoing edge in a 1-ring neighborhood of a vertex, and average pooling aggregates per-half-flap outputs into a per-vertex feature. **EDGE** step only considers per-vertex features at two endpoints of a subdivided edge to compute the feature of the inserted vertex. Thus, it simply applies half-flap operator for each directions of the edge and again uses average pooling to get the vertex feature.

The final critical element of our architecture design is the representation for the input and output. As mentioned before, we use local differential quantities to ensure invariance to rigid transformation. The input features for the half-flaps used in **INITIALIZATION** step by module  $\mathcal{I}$  consist of **three edge vectors (originating at the source vertex of half-flap)** and **differential coordinates of each vertex**, as illustrated in Fig. 19, top. The vector of differential coordinates stores **the discrete curvature information** and is defined as the difference between the absolute coordinates of a vertex and the average of its immediate neighbors in the mesh [Sorkine 2005]. To achieve rotation invariance we represent our differential quantities in the local frame of each half-flap (see the inset), where we treat the half-edge direction as the x-axis, the edge normal computed via averaging the two adjacent face normals as the z-axis, and the cross product of the previous two axes becomes



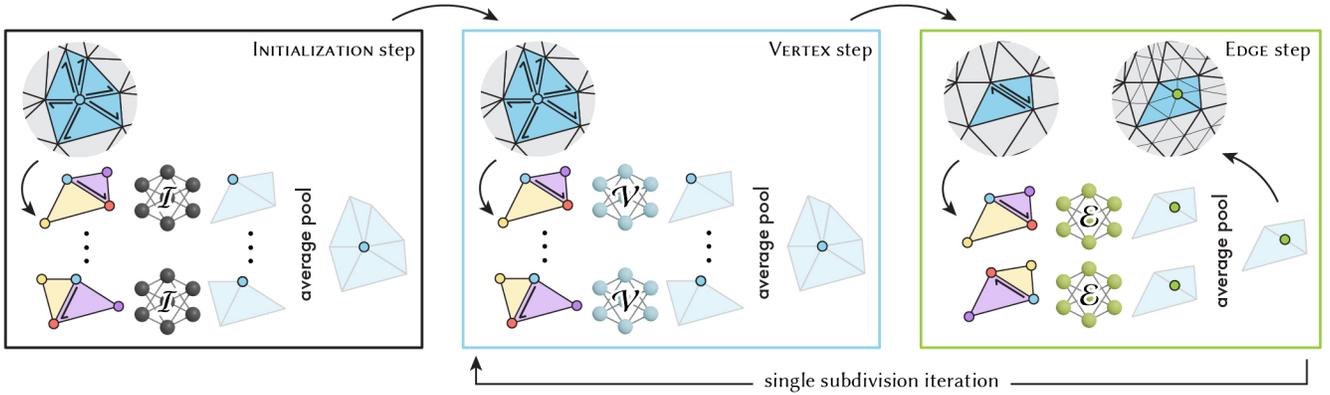


Fig. 18. Our neural subdivision consists of three sequential steps: **INITIALIZATION**, **VERTEX**, and **EDGE**, with three network modules:  $\mathcal{I}$ ,  $\mathcal{V}$ , and  $\mathcal{E}$  for each step respectively. In both **INITIALIZATION** and **VERTEX** steps, we apply  $\mathcal{V}$  and  $\mathcal{E}$  for the half-flaps of all the outgoing edges of a vertex, and use average pooling to combine the output features back to the center vertex (blue). In the **EDGE** step, we apply  $\mathcal{E}$  to both half-flaps of an undirected edge and use average pooling to map the output features to the center vertex (green) of the edge.

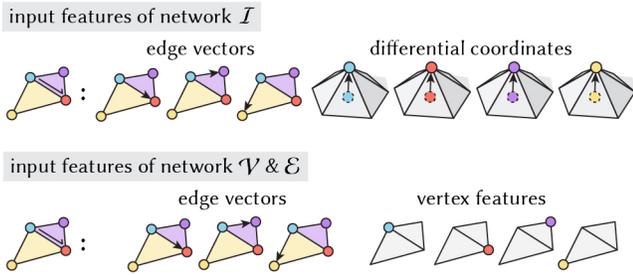


Fig. 19. The input feature to module  $\mathcal{I}$  consists of three edge vectors from the source vertex (blue) and vectors of the differential coordinates for the four vertices. The input features to module  $\mathcal{V}$  and  $\mathcal{E}$  are three edge vectors with per-vertex high-dimensional features from the previous steps.

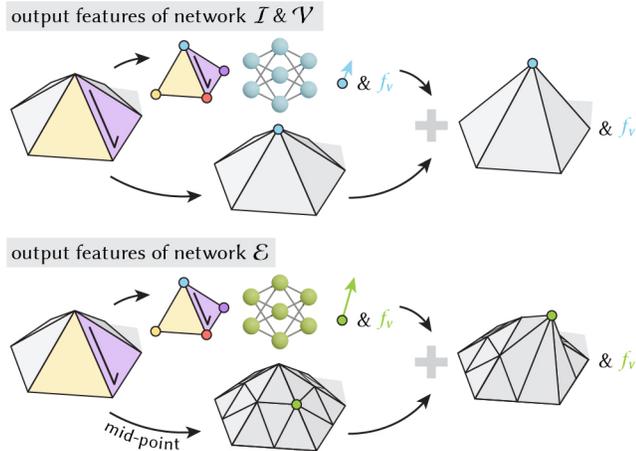


Fig. 20. The outputs of modules  $\mathcal{I}$  and  $\mathcal{V}$  are the displacement vector from the starting vertex and a learned feature vector  $f_v$  stored at the source vertex (blue). The outputs of the module  $\mathcal{E}$  are the displacement from the edge mid-point (green) and the feature  $f_v$  stored at the mid-point.

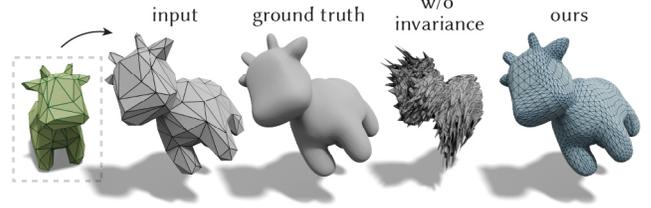


Fig. 21. We use differential quantities stored in the local frames as our inputs and outputs. This design makes our network invariant to rigid motions and significantly boosts the quality compared to an approach without invariance.

our y-axis. The input to half-flap operators used in **VERTEX** and **EDGE** steps is similar (Fig. 19, bottom), where we use **edge vectors** and **per-vertex high-dimensional learned features** (either produced by **INITIALIZATION** step or by previous subdivision iteration). The output of half-flaps used in **VERTEX** and **EDGE** steps includes high-dimensional learned features and differential quantities that can be used to reconstruct the vertex position. For the latter we use the vertex displacement vector from the mid-point subdivided mesh (see Fig. 20) in the local coordinate system of the half-flap. For the **INITIALIZATION** and the **VERTEX** networks, **the predicted displacements live on the vertices**; for the **EDGE** network, **the predicted displacements live on the edge midpoints**. In our experiments, we notice there is **no difference between predicting from the mid-point subdivided surface or other subdivision surfaces** (see App. H), so we choose mid-point subdivision for simplicity. We estimate global coordinates of vertices after each step to visualize intermediate levels of subdivision and compute the loss function, and convert global coordinates to local differential per-vertex quantities before each step to ensure that each network only observes **translation- and rotation- invariant representations**.

Fig. 21 illustrates that invariant representation is critical to the quality of results. We demonstrate that even when trained on an identical true shape, a slight rigid motion of that shape renders learned weights completely inapplicable at inference time. We also

Table 1. Hyperparameters of our sub-networks. All networks are fully-connected multi-layer perceptrons with two hidden layers.

	network $\mathcal{I}$	network $\mathcal{V}$	network $\mathcal{E}$
$f_{in}$	$3 \cdot 3 + 4 \cdot 3$	$3 \cdot 3 + 4 \cdot 32$	$3 \cdot 3 + 4 \cdot 32$
$f_{c_1}$	32	32	32
$f_{c_2}$	32	32	32
$f_{out}$	$3 + 29$	$3 + 29$	$3 + 29$

observe that incorporating the differential coordinates as part of the input features makes the training converge faster (see App. H). Thanks to our local half-flap operators and invariant representations we can train our architecture even with shallow 2-layer MLPs (see Table 1 for network hyper-parameters). We further evaluate other design decisions and conclude that details such as whether to predict displacements from the mid-point or the Loop subdivision, whether to recursively apply the module  $\mathcal{V}$ , whether to measure loss across all levels, and whether to use input features proposed in [Hanocka et al. 2019] offer small improvements to the convergence (see App. H for details).

We implemented our network in PyTorch [Paszke et al. 2019]. We use ReLU activation [Nair and Hinton 2010], and the ADAM optimizer [Kingma and Ba 2015] with learning rate 0.002.

## 6 EVALUATIONS

We evaluate our neural subdivision with a range of results of increasing complexity. We start by showing that we can generalize to isometric deformations, non-isometric deformations, shapes from different classes, and shapes from different types of discretizations. We summarize the details of our experiments in App. F.

In practice, modelers often manipulate the coarse subdivision cage of a character into different poses, and then apply the subdivision operator. This scenario implies that being able to train on one single pose and generalize to unseen poses is important for character animation. In Fig. 22, we train on a single pose (in green) and show that our network can generalize to unseen poses under (approximately) isometric deformations.

In addition to poses, in Fig. 23 we mimic the real scenario to manually change the coarse cage and show that the learned subdivision can also generalize to non-isometric deformations.

Subdivision operators are often used to create novel 3D content, which implies the importance of generalizing to totally different shapes. In Fig. 24 we show that even when trained on only a single shape (green), our network is able to generalize to many other shapes (blue). We also show that our network trained on classic Loop subdivision sequences is able to reproduce Loop subdivision on unseen shapes (App. G).

We further evaluate neural subdivision on shape discretizations created in a totally independent way. In Fig. 25 we obtain coarse shapes created by artists, instead of from edge collapses, and show that neural subdivision can still generalize well.

The ability to generalize even when trained on a single shape gives us the opportunity to do stylized subdivisions. In Fig. 26 our neural subdivision operators are aware of the “style” of the training shape and are able to create different results from the same coarse

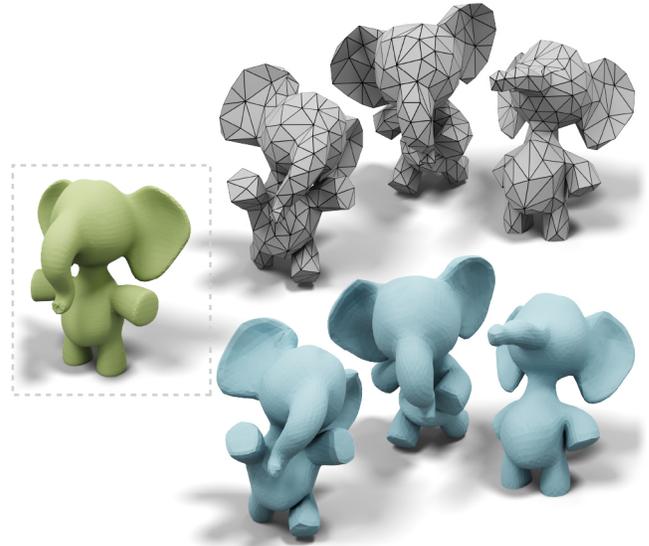


Fig. 22. We train our network on a single pose (green) and the network is able to generalize to unseen poses (blue).

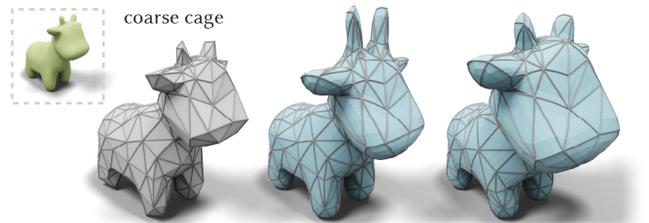


Fig. 23. We mimic the modeling scenario by applying non-isometric deformations to the coarse cage (gray). Our subdivision network is able to generalize to unseen non-isometric deformations.

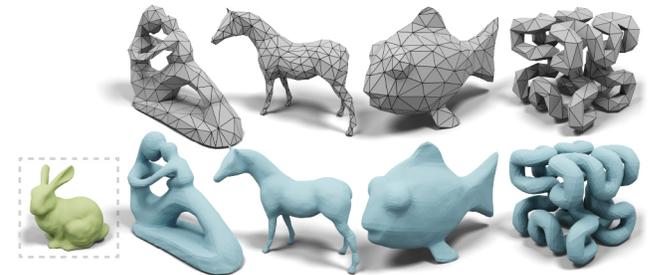


Fig. 24. Even when trained on only a single shape (green bunny), our network can generalize to subdividing different geometries (blue). ©Hilbert Cube by tbusser (right) under CC BY-NC.

geometry. In Fig. 27, we show different results when trained on a smooth organic shape vs a man-made object with sharp contours.

To quantitatively analyze how our network generalizes to unseen shapes, we take the TOSCA dataset [Bronstein et al. 2009] which contains 80 shapes with 9 categories to perform quantitative analysis. For the top table of Table 2, we train on a single category (*Centaur*) and test on the remaining categories. Our test shapes are generated

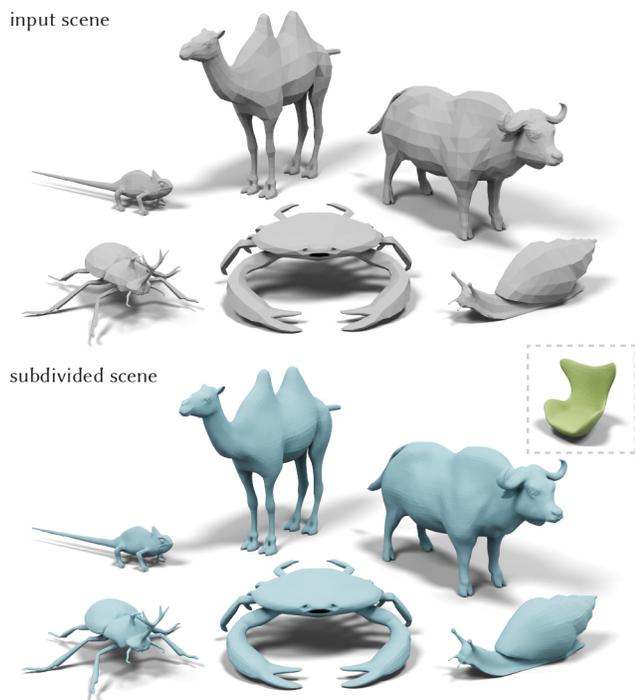


Fig. 25. In addition to subdividing meshes constructed via decimation, our network can also generalize to subdivide meshes created by artists.

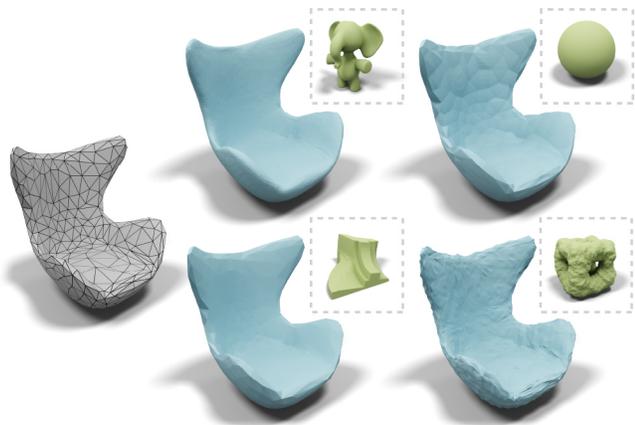


Fig. 26. Using different shapes in training leads to stylized subdivision results (blue) biased towards the training shapes (green). ©Egg Chair by TeamTeamUSA (left) under CC BY.

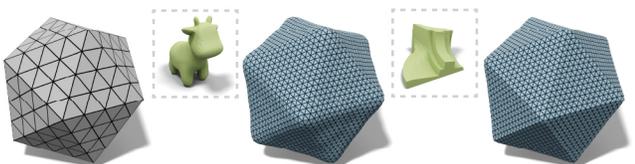


Fig. 27. Training on a smooth shape leads to a smoother subdivision result (middle). Training on a man-made object can preserve the sharp creases (right).

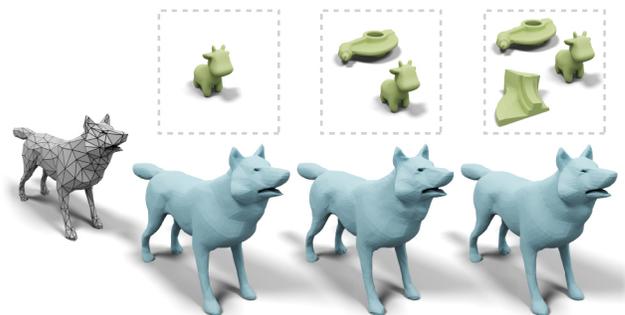


Fig. 28. We train our subdivision network on a mixture of organic and non-organic shapes. We observe that training on more objects does not significantly change visual quality in this case.

Table 2. We train on a single category, *Centaur* (top table), and three categories, *Centaur*, *David*, *Horse* (bottom table), separately, and evaluate by subdividing the rest of the TOSCA shapes. The results indicate that neural subdivision outperforms classic Loop subdivision [Loop 1987] and modified butterfly subdivision [Zorin et al. 1996] on two popular metrics: Hausdorff distance  $\mathbb{H}$ , and mean surface distance  $\mathbb{M}$  computed via METRO [Cignoni et al. 1998].

Category	$\mathbb{H}_{\text{loop}}$	$\mathbb{H}_{\text{m.b.}}$	$\mathbb{H}_{\text{ours}}$	$\mathbb{M}_{\text{loop}}$	$\mathbb{M}_{\text{m.b.}}$	$\mathbb{M}_{\text{ours}}$
Cat	2.75	2.17	<b>2.08</b>	0.73	0.21	<b>0.17</b>
David	2.95	2.13	<b>1.83</b>	0.88	0.27	<b>0.20</b>
Dog	3.26	2.32	<b>2.11</b>	0.84	0.31	<b>0.26</b>
Gorilla	4.53	3.17	<b>2.56</b>	1.27	0.48	<b>0.36</b>
Horse	5.87	4.53	<b>4.04</b>	1.51	0.50	<b>0.45</b>
Michael	3.88	2.71	<b>2.24</b>	1.12	0.38	<b>0.28</b>
Victoria	4.25	3.01	<b>2.36</b>	1.12	0.39	<b>0.30</b>
Wolf	2.83	1.74	<b>1.63</b>	0.69	0.23	<b>0.21</b>

Category	$\mathbb{H}_{\text{loop}}$	$\mathbb{H}_{\text{m.b.}}$	$\mathbb{H}_{\text{ours}}$	$\mathbb{M}_{\text{loop}}$	$\mathbb{M}_{\text{m.b.}}$	$\mathbb{M}_{\text{ours}}$
Cat	2.75	2.17	<b>2.09</b>	0.73	0.21	<b>0.16</b>
Dog	3.26	2.32	<b>2.12</b>	0.84	0.31	<b>0.25</b>
Gorilla	4.53	3.17	<b>2.89</b>	1.27	0.48	<b>0.34</b>
Michael	3.88	2.71	<b>2.15</b>	1.12	0.38	<b>0.27</b>
Victoria	4.25	3.01	<b>2.49</b>	1.12	0.39	<b>0.28</b>
Wolf	2.83	1.74	<b>1.65</b>	0.69	0.23	<b>0.20</b>

by coarsening source meshes with QSLIM down to 350-450 vertices. We measure the error between the two-level subdivided mesh and the original shape using Hausdorff distance, as well as mean surface distance computed by the METRO [Cignoni et al. 1998]. Our method consistently produces smaller errors compared to the classic Loop [Loop 1987] and modified butterfly [Zorin et al. 1996] subdivisions.

We further evaluate our method when trained on multiple shapes and categories. In Fig. 28, we train the network on an increasing number of objects and observe that the results are visually similar. But our quantitative analysis in the bottom table of Table 2 shows that training on more categories (*Centaur*, *David*, *Horse*) can slightly reduce the error.

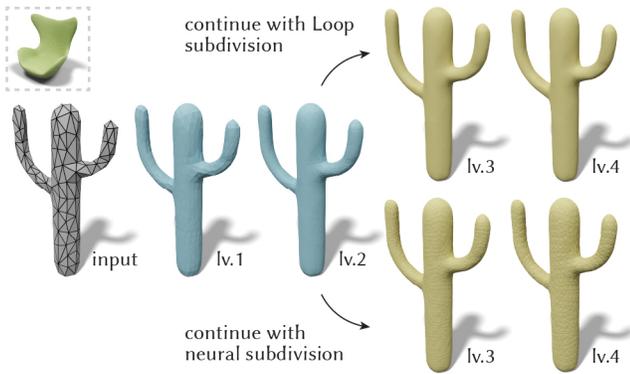


Fig. 29. Since our method induces a non-linear subdivision, there is no guarantee for the existence of a limit surface (bottom). An alternative is to apply neural subdivision at the trained levels, and continue with classic subdivision (top) to ensure a smooth limit surface.

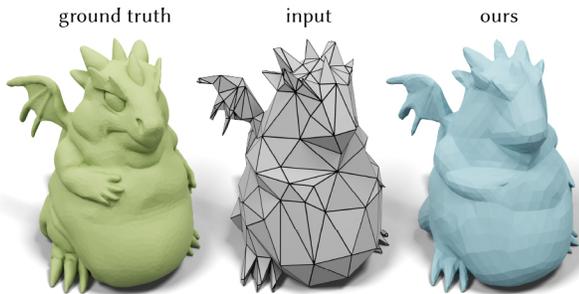


Fig. 30. Our approach is based on local geometry, and thus fails to hallucinate semantic features. ©Bratty Dragon by Splotchy Ink under CC BY.

## 7 LIMITATIONS & FUTURE WORK

Extending the neural subdivision framework to quadrilateral meshes and surface with boundaries would be closer to real-world modeling scenarios. Making neural subdivision scale-invariant and converge to a limit surface (see Fig. 29) are also desirable in practice. Incorporating global information in the training could help the network hallucinate semantic features (see Fig. 30). Applying architectures (e.g., Recurrent Neural Net) that are more suitable for sequence predictions could help the network to harness information from a wider neighborhood and to dive to a deeper subdivision level. Training on data that contain a wide range of triangle aspect ratios and curvature information could further improve the robustness of the network. Since our data-generation algorithm is extremely efficient, it could be naturally used in an online-learning setting, where our algorithm constantly draws new randomly-coarsened meshes on-the-fly. This can be extremely useful in, e.g., a GAN setting. As a first step towards neural subdivision, we showed reconstruction of fine meshes from coarse ones. Fully-fledged super-resolution, detail hallucination, and surface stylization are interesting next steps. All of these questions provide interesting topics for the future research on neural subdivision.

## ACKNOWLEDGMENTS

Our research is funded in part by New Frontiers of Research Fund (NFRFE-201), the Ontario Early Research Award program, NSERC Discovery (RGPIN2017-05235, RGPAS-2017-507938), the Canada Research Chairs Program, the Fields Centre for Quantitative Analysis and Modelling and gifts by Adobe Systems, Autodesk and MESH Inc. We thank members of Dynamic Graphics Project at the University of Toronto; Thomas Davies, Oded Stein, Michael Tao, and Jackson Wang for early discussions; Rahul Arora, Seungbae Bang, Jiannan Li, Abhishek Madan, and Silvia Sellán for experiments and generating test data; Honglin Chen, Eitan Grinspun, and Sarah Kushner for proofreading. We thank Mirela Ben-Chen for the enlightening advice on the experiments and the writing; Yifan Wang for running comparisons; and Ahmad Nasikun for evaluations. We obtained our test models from THING10K [Zhou and Jacobson 2016] and we thank all the artists for sharing a rich variety of 3D models. We especially thank John Hancock for the IT support which helped us smoothly conduct this research.

## REFERENCES

- Noam Aigerman, Roi Poranne, and Yaron Lipman. 2014. Lifted bijections for low distortion surface mappings. *ACM Transactions on Graphics (TOG)* 33, 4 (2014), 69.
- Noam Aigerman, Roi Poranne, and Yaron Lipman. 2015. Seamless surface mappings. *ACM Transactions on Graphics (TOG)* 34, 4 (2015), 72.
- Harry G. Barrow, Jay M. Tenenbaum, Robert C. Bolles, and Helen C. Wolf. 1977. Parametric Correspondence and Chamfer Matching: Two New Techniques for Image Matching. In *Proceedings of the 5th International Joint Conference on Artificial Intelligence*. Cambridge, MA, USA, August 22-25, 1977. Raj Reddy (Ed.). William Kaufmann, 659–663.
- Alexander I. Bobenko, Helmut Pottmann, and Thilo Rörig. 2020. Multi-Nets. Classification of Discrete and Smooth Surfaces with Characteristic Properties on Arbitrary Parameter Rectangles. *Discret. Comput. Geom.* 63, 3 (2020), 624–655.
- Alexander M. Bronstein, Michael M. Bronstein, and Ron Kimmel. 2009. *Numerical Geometry of Non-Rigid Shapes*. Springer.
- Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst. 2017. Geometric Deep Learning: Going beyond Euclidean data. *IEEE Signal Process. Mag.* 34, 4 (2017), 18–42.
- E. Catmull and J. Clark. 1998. *Recursively Generated B-Spline Surfaces on Arbitrary Topological Meshes*. Association for Computing Machinery, New York, NY, USA, 183–188.
- Siddhartha Chaudhuri, Daniel Ritchie, Jiajun Wu, Kai Xu, and Hao (Richard) Zhang. 2020. Learning to Generate 3D Structures. In *Eurographics State-of-the-Art Report (STAR)*.
- Paolo Cignoni, Claudio Rocchini, and Roberto Scopigno. 1998. Metro: measuring error on simplified surfaces. In *Computer graphics forum*, Vol. 17. Wiley Online Library, 167–174.
- Jonathan Cohen, Marc Olano, and Dinesh Manocha. 1998. Appearance-Preserving Simplification. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '98)*. Association for Computing Machinery, New York, NY, USA, 115–122.
- Jonathan D. Cohen, Dinesh Manocha, and Marc Olano. 1997. Simplifying polygonal models using successive mappings. In *IEEE Visualization '97, Proceedings, Phoenix, AZ, USA, October 19-24, 1997*. IEEE Computer Society and ACM, 395–402.
- Jonathan D. Cohen, Dinesh Manocha, and Marc Olano. 2003. Successive Mappings: An Approach to Polygonal Mesh Simplification with Guaranteed Error Bounds. *Int. J. Comput. Geometry Appl.* 13, 1 (2003), 61.
- Angela Dai and Matthias Nießner. 2019. Scan2Mesh: From Unstructured Range Scans to 3D Meshes. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, 5574–5583.
- Tony DeRose, Michael Kass, and Tien Truong. 1998. Subdivision Surfaces in Character Animation. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1998, Orlando, FL, USA, July 19-24, 1998*, Steve Cunningham, Walt Bransford, and Michael F. Cohen (Eds.). ACM, 85–94.
- Tamal K Dey, Herbert Edelsbrunner, Sumanta Guha, and Dmitry V Nekhayev. 1999. Topology preserving edge contraction. *Publ. Inst. Math.(Beograd)(NS)* 66, 80 (1999), 23–45.
- Daniel Doo. 1978. A subdivision algorithm for smoothing down irregularly shaped polyhedrons. *Computer Aided Design* (1978), 157–165.

- D. Doo and M. Sabin. 1998. *Behaviour of Recursive Division Surfaces near Extraordinary Points*. Association for Computing Machinery, New York, NY, USA, 177–181.
- Nira Dyn, David Levine, and John A. Gregory. 1990. A butterfly subdivision scheme for surface interpolation with tension control. *ACM Trans. Graph.* 9, 2 (1990), 160–169.
- Haoqiang Fan, Hao Su, and Leonidas J. Guibas. 2017. A Point Set Generation Network for 3D Object Reconstruction from a Single Image. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21–26, 2017*. IEEE Computer Society, 2463–2471.
- Michael S. Floater and Charles A. Micchelli. 1997. Nonlinear Stationary Subdivision. *Journal of Approximation Theory* (1997).
- Michael Garland and Paul S. Heckbert. 1997. Surface simplification using quadric error metrics. In *Proceedings of the 24th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1997, Los Angeles, CA, USA, August 3–8, 1997*, G. Scott Owen, Turner Whitted, and Barbara Mones-Hattal (Eds.). ACM, 209–216.
- Thibault Groueix, Matthew Fisher, Vladimir G. Kim, Bryan C. Russell, and Mathieu Aubry. 2018a. 3D-CODED: 3D Correspondences by Deep Deformation. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part II (Lecture Notes in Computer Science)*, Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss (Eds.), Vol. 11206. Springer, 235–251.
- Thibault Groueix, Matthew Fisher, Vladimir G. Kim, Bryan C. Russell, and Mathieu Aubry. 2018b. A Papier-Mâché Approach to Learning 3D Surface Generation. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18–22, 2018*. IEEE Computer Society, 216–224.
- Igor Guskov, Andrei Khodakovskiy, Peter Schröder, and Wim Sweldens. 2002. Hybrid meshes: multiresolution using regular and irregular refinement. In *Proceedings of the 18th Annual Symposium on Computational Geometry, Barcelona, Spain, June 5–7, 2002*, Ferran Hurtado, Vera Sacristán, Chandrajit Bajaj, and Subhash Suri (Eds.). ACM, 264–272.
- Igor Guskov, Kiril Vidimce, Wim Sweldens, and Peter Schröder. 2000. Normal meshes. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2000, New Orleans, LA, USA, July 23–28, 2000*, Judith R. Brown and Kurt Akeley (Eds.). ACM, 95–102.
- Rana Hanocka, Amir Hertz, Noa Fish, Raja Giryes, Shachar Fleishman, and Daniel Cohen-Or. 2019. MeshCNN: A Network with an Edge. *ACM Transactions on Graphics (TOG)* 38, 4 (2019), 90.
- Hugues Hoppe, Tony DeRose, Tom Duchamp, Mark A. Halstead, Hubert Jin, John Alan McDonald, Jean Schweitzer, and Werner Stuetzle. 1994. Piecewise smooth surface reconstruction. In *Proceedings of the 21th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1994, Orlando, FL, USA, July 24–29, 1994*, Dino Schweitzer, Andrew S. Glassner, and Mike Keeler (Eds.). ACM, 295–302.
- Hugues Hoppe, Tony DeRose, Tom Duchamp, John Alan McDonald, and Werner Stuetzle. 1993. Mesh optimization. In *Proceedings of the 20th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1993, Anaheim, CA, USA, August 2–6, 1993*, Mary C. Whitton (Ed.). ACM, 19–26.
- Krishna Murthy J., Edward Smith, Jean-Francois Lafleche, Clement Fuji Tsang, Artem Rozantsev, Wenzheng Chen, Tommy Xiang, Rev Lebareddan, and Sanja Fidler. 2019. Kaolin: A PyTorch Library for Accelerating 3D Deep Learning Research. *arXiv:1911.05063* (2019).
- Kestutis Karčiauskas and Jörg Peters. 2018. A New Class of Guided C2 Subdivision Surfaces Combining Good Shape with Nested Refinement. *Comput. Graph. Forum* 37, 6 (2018), 84–95.
- Ladislav Kavan, Dan Gerszewski, Adam W. Bargteil, and Peter-Pike Sloan. 2011. Physics-Inspired Upsampling for Cloth Simulation in Games. *ACM Trans. Graph.* 30, 4, Article Article 93 (July 2011), 10 pages.
- Michael M. Kazhdan and Hugues Hoppe. 2013. Screened poisson surface reconstruction. *ACM Trans. Graph.* 32, 3 (2013), 29:1–29:13.
- Andrei Khodakovskiy, Nathan Litke, and Peter Schröder. 2003. Globally Smooth Parameterizations with Low Distortion. *ACM Trans. Graph.* 22, 3 (July 2003), 350–357. <https://doi.org/10.1145/882262.882275>
- Diederik P. Kingma and Jimmy Ba. 2015. Adam: A Method for Stochastic Optimization. In *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7–9, 2015, Conference Track Proceedings*, Yoshua Bengio and Yann LeCun (Eds.).
- Leif Kobbelt. 1996. Interpolatory Subdivision on Open Quadrilateral Nets with Arbitrary Topology. *Comput. Graph. Forum* 15, 3 (1996), 409–420.
- Leif Kobbelt. 2000. 3-subdivision. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 2000, New Orleans, LA, USA, July 23–28, 2000*, Judith R. Brown and Kurt Akeley (Eds.). ACM, 103–112.
- Ilya Kostrikov, Zhongshi Jiang, Daniele Panozzo, Denis Zorin, and Joan Bruna. 2018. Surface Networks. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18–22, 2018*. IEEE Computer Society, 2540–2548.
- Vladislav Kraevoy and Alla Sheffer. 2004. Cross-parameterization and compatible remeshing of 3D models. *ACM Transactions on Graphics (TOG)* 23, 3 (2004), 861–869.
- Aaron W. F. Lee, Wim Sweldens, Peter Schröder, Lawrence C. Cowsar, and David P. Dobkin. 1998. MAPS: Multiresolution Adaptive Parameterization of Surfaces. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1998, Orlando, FL, USA, July 19–24, 1998*, Steve Cunningham, Walt Bransford, and Michael F. Cohen (Eds.). ACM, 95–104.
- Ruihui Li, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. 2019. PU-GAN: A Point Cloud Upsampling Adversarial Network. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 – November 2, 2019*. IEEE, 7202–7211.
- Songrun Liu, Zachary Ferguson, Alec Jacobson, and Yotam I. Gingold. 2017. Seamless: seam erasure and seam-aware decoupling of shape from mesh resolution. *ACM Trans. Graph.* 36, 6 (2017), 216:1–216:15.
- Yang Liu, Helmut Pottmann, Johannes Wallner, Yong-Liang Yang, and Wenping Wang. 2006. Geometric modeling with conical meshes and developable surfaces. In *ACM transactions on graphics (TOG)*, Vol. 25. ACM, 681–689.
- Charles Loop. 1987. Smooth subdivision surfaces based on triangles. *Master's thesis, University of Utah, Department of Mathematics* (1987).
- Haggai Maron, Meirav Galun, Noam Aigerman, Miri Trope, Nadav Dym, Ersin Yumer, Vladimir G. Kim, and Yaron Lipman. 2017. Convolutional neural networks on surfaces via seamless toric covers. *ACM Trans. Graph.* 36, 4 (2017), 71:1–71:10.
- Jonathan Masci, Davide Boscaini, Michael M. Bronstein, and Pierre Vandergheynst. 2015. Geodesic Convolutional Neural Networks on Riemannian Manifolds. In *2015 IEEE International Conference on Computer Vision Workshop, ICCV Workshops 2015, Santiago, Chile, December 7–13, 2015*. IEEE Computer Society, 832–840.
- Patrick Mullen, Yiyang Tong, Pierre Alliez, and Mathieu Desbrun. 2008. Spectral Conformal Parameterization. *Comput. Graph. Forum* 27, 5 (2008), 1487–1494.
- Vinod Nair and Geoffrey E. Hinton. 2010. Rectified Linear Units Improve Restricted Boltzmann Machines. In *Proceedings of the 27th International Conference on Machine Learning (ICML-10), June 21–24, 2010, Haifa, Israel*, Johannes Fürnkranz and Thorsten Joachims (Eds.). Omnipress, 807–814.
- Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Köpf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library. In *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, 8–14 December 2019, Vancouver, BC, Canada*, Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett (Eds.). 8024–8035.
- Adrien Poulenard and Maks Ovsjanikov. 2018. Multi-directional geodesic neural networks via equivariant convolution. *ACM Trans. Graph.* 37, 6 (2018), 236:1–236:14.
- Emil Praun, Wim Sweldens, and Peter Schröder. 2001. Consistent Mesh Parameterizations. In *Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '01)*. Association for Computing Machinery, New York, NY, USA, 179–184.
- Reinhold Preiner, Tamy Boubekeur, and Michael Wimmer. 2019. Gaussian-product subdivision surfaces. *ACM Trans. Graph.* 38, 4 (2019), 35:1–35:11.
- Michael Rabinovich, Tim Hoffmann, and Olga Sorkine-Hornung. 2018. The shape space of discrete orthogonal geodesic nets. *ACM Trans. Graph.* 37, 6 (2018), 228:1–228:17.
- Anurag Ranjan, Timo Bolkart, Souhik Sanyal, and Michael J. Black. 2018. Generating 3D Faces Using Convolutional Mesh Autoencoders. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part III (Lecture Notes in Computer Science)*, Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss (Eds.), Vol. 11207. Springer, 725–741.
- Malcolm Sabin and Neil Dodgson. 2004. A Circle-Preserving Variant of the Four-Point Subdivision Scheme. *Mathematical Methods for Curves and Surfaces: Tromsø 2004* (01 2004).
- Scott Schaefer, E. Vouga, and Ron Goldman. 2008. Nonlinear subdivision through nonlinear averaging. *Comput. Aided Geom. Des.* 25, 3 (2008), 162–180.
- John Schreiner, Arul Asirvatham, Emil Praun, and Hugues Hoppe. 2004. Inter-surface mapping. *ACM Trans. Graph.* 23, 3 (2004), 870–877.
- Vincent Sitzmann, Justus Thies, Felix Heide, Matthias Nießner, Gordon Wetzstein, and Michael Zollhöfer. 2019. DeepVoxels: Learning Persistent 3D Feature Embeddings. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16–20, 2019*. Computer Vision Foundation / IEEE, 2437–2446.
- Olga Sorkine. 2005. Laplacian Mesh Processing. In *Eurographics 2005 - State of the Art Reports, Dublin, Ireland, August 29 - September 2, 2005*, Yiorgos Chrysanthou and Marcus A. Magnor (Eds.). Eurographics Association, 53–70.
- Jos Stam. 1998. Exact Evaluation of Catmull-Clark Subdivision Surfaces at Arbitrary Parameter Values. In *Proceedings of the 25th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH 1998, Orlando, FL, USA, July 19–24, 1998*, Steve Cunningham, Walt Bransford, and Michael F. Cohen (Eds.). ACM, 395–404.
- Qingyang Tan, Lin Gao, Yu-Kun Lai, and Shihong Xia. 2018. Variational Autoencoders for Deforming 3D Mesh Models. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18–22, 2018*. IEEE Computer Society, 5841–5850.

- Chengcheng Tang, Xiang Sun, Alexandra Gomes, Johannes Wallner, and Helmut Pottmann. 2014. Form-Finding with Polyhedral Meshes Made Simple. *ACM Trans. Graph.* 33, 4, Article Article 70 (July 2014), 9 pages.
- Maxim Tatarchenko, Stephan R. Richter, René Ranftl, Zhuwen Li, Vladlen Koltun, and Thomas Brox. 2019. What Do Single-View 3D Reconstruction Networks Learn?. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, 3405–3414.
- Robert F. Tobler, Stefan Maierhofer, and Alexander Wilkie. 2002a. Mesh-Based Parametrized L-Systems and Generalized Subdivision for Generating Complex Geometry. *International Journal of Shape Modeling* 8, 2 (2002), 173–191.
- Robert F. Tobler, Stefan Maierhofer, and Alexander Wilkie. 2002b. A Multiresolution Mesh Generation Approach for Procedural Definition of Complex Geometry. In *2002 International Conference on Shape Modeling and Applications (SMI 2002), 17-22 May 2002, Banff, Alberta, Canada*. IEEE Computer Society, 35–42.
- Oliver Van Kaick, Hao Zhang, Ghassan Hamarneh, and Daniel Cohen-Or. 2011. A survey on shape correspondence. In *Computer Graphics Forum*, Vol. 30. Wiley Online Library, 1681–1707.
- Amir Vaxman, Christian Müller, and Ofir Weber. 2018. Canonical Möbius subdivision. *ACM Trans. Graph.* 37, 6 (2018), 227:1–227:15.
- Luiz Velho, Ken Perlin, Henning Biermann, and Lexing Ying. 2002. Algorithmic shape modeling with subdivision surfaces. *Comput. Graph.* 26, 6 (2002), 865–875.
- Matthias Vestner, Roei Litman, Emanuele Rodolà, Alex Bronstein, and Daniel Cremers. 2017. Product manifold filter: Non-rigid shape correspondence via kernel density estimation in the product space. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 3327–3336.
- Nanyang Wang, Yinda Zhang, Zhuwen Li, Yanwei Fu, Wei Liu, and Yu-Gang Jiang. 2018. Pixel2Mesh: Generating 3D Mesh Models from Single RGB Images. In *Computer Vision - ECCV 2018 - 15th European Conference, Munich, Germany, September 8-14, 2018, Proceedings, Part XI (Lecture Notes in Computer Science)*, Vittorio Ferrari, Martial Hebert, Cristian Sminchisescu, and Yair Weiss (Eds.), Vol. 11215. Springer, 55–71.
- Yu Wang, Vladimir G. Kim, Michael Bronstein, and Justin Solomon. 2019a. Learning Geometric Operators on Meshes. *ICLR Workshop on Representation Learning on Graphs and Manifolds* (2019).
- Yifan Wang, Shihao Wu, Hui Huang, Daniel Cohen-Or, and Olga Sorkine-Hornung. 2019b. Patch-Based Progressive 3D Point Set Upsampling. In *IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2019, Long Beach, CA, USA, June 16-20, 2019*. Computer Vision Foundation / IEEE, 5958–5967.
- Chao Wen, Yinda Zhang, Zhuwen Li, and Yanwei Fu. 2019. Pixel2Mesh++: Multi-View 3D Mesh Generation via Deformation. In *2019 IEEE/CVF International Conference on Computer Vision, ICCV 2019, Seoul, Korea (South), October 27 - November 2, 2019*. IEEE, 1042–1051.
- Wang Yifan, Noam Aigerman, Vladimir G. Kim, Siddhartha Chaudhuri, and Olga Sorkine-Hornung. 2020. Neural Cages for Detail-Preserving 3D Deformations. In *CVPR*.
- Lequan Yu, Xianzhi Li, Chi-Wing Fu, Daniel Cohen-Or, and Pheng-Ann Heng. 2018. PU-Net: Point Cloud Upsampling Network. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*. IEEE Computer Society, 2790–2799.
- Qingnan Zhou and Alec Jacobson. 2016. Thing10K: A Dataset of 10,000 3D-Printing Models. *arXiv preprint arXiv:1605.04797* (2016). <https://ten-thousand-models.appspot.com>
- Denis Zorin. 2007. Subdivision on arbitrary meshes: algorithms and theory. In *Mathematics and Computation in Imaging Science and Information Processing*. World Scientific, 1–46.
- Denis Zorin, Peter Schröder, T De Rose, L Kobbelt, A Levin, and W Sweldens. 2000. Subdivision for modeling and animation. *SIGGRAPH Course Notes* (2000).
- Denis Zorin, Peter Schröder, and Wim Sweldens. 1996. Interpolating Subdivision for Meshes with Arbitrary Topology. In *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques (SIGGRAPH '96)*. Association for Computing Machinery, New York, NY, USA, 189–192.

## A IMPLEMENTATION OF POINT CLOUD UPSAMPLING

An alternative way to upsample a mesh is to first convert the mesh into point cloud via sampling over the surface, run point cloud upsampling algorithms, and then perform a surface reconstruction to convert the upsampled point cloud back to a mesh. However, this procedure is expensive to incorporate into the interactive graphics pipeline, fails to produce surfaces with different levels of detail (see Fig. 2), and it fails to preserve textures (see Fig. 3). In addition, many non-trivial design decisions such as the number of samples to use and how to sample the surface would influence the quality of

the results. For example in Fig. 5, we first sample 5000 points with uniform and farthest point sampling, followed by the method of Wang et al. [2019b] pre-trained on statues to upsample the point cloud by 16×, and then use the screened poisson reconstruction [Kazhdan and Hoppe 2013] to reconstruct the surface. In the figure we show that different sampling methods lead to different results. The lack of connectivity information also results in some surface artifacts.

## B IMPLEMENTATION OF SUCCESSIVE SELF-PARAMETERIZATION

Incorporating successive self-parameterization only requires adding two additional local conformal parameterizations to the edge collapse algorithm of choice. Suppose we want to collapse an edge  $(j, k)$ . We first flatten the edge’s 1-ring  $\mathcal{N}(j, k)$ , then we collapse the edge, then we perform another conformal flattening on the 1-ring  $\mathcal{N}(i)$  of the newly inserted vertex  $i$  after the collapse, with the boundary held to place from the previous flattening. This yields a bijective map with small computational cost because each flattening only involves a 1-ring (assuming the vertex valence is bounded).

## C CRITERIA FOR COLLAPSIBLE EDGES

During edge collapses, many issues such as flipped faces and non-manifold edges may appear. Resolving these issues is crucial to the robustness of successive self-parameterization (see Fig. 16). We summarize our criteria for checking the validity of an edge collapse. If invalid, we simply avoid collapsing the edge at that iteration.

*Euclidean face flips.* Certain faces in the Euclidean space may suffer from normal flips after an edge collapse. To prevent flipped faces, we simply compare the unit face normal  $\hat{n}$  of each neighboring face  $f_i$  before and after the collapse

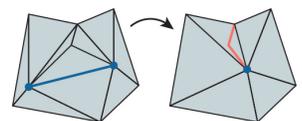
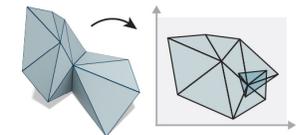
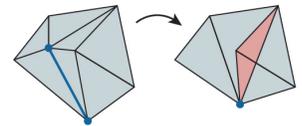
$$\hat{n}_{f_i}^{\text{before}} \cdot \hat{n}_{f_i}^{\text{after}} > \delta. \quad (2)$$

Our default  $\delta = 0.2$  which is sufficient to avoid face flips in all our experiments.

*UV face flips.* Flipped faces may also appear in the UV space due to both the conformal flattening and the edge collapse. We simply check whether the signed area of each UV face is positive before and after collapses to prevent having UV face flips.

*Overlapped UV faces.* Even if all the UV faces are oriented correctly, some of the faces may still overlap with each other depending on the flattening algorithm in use. We check whether the total angle sum of each interior vertex is  $2\pi$  to determine the validity of a collapse.

*Non-manifold edges.* To prevent the appearance of non-manifold edges, we must check the *link condition* [Dey et al. 1999; Hoppe et al. 1993]. Briefly, the link condition



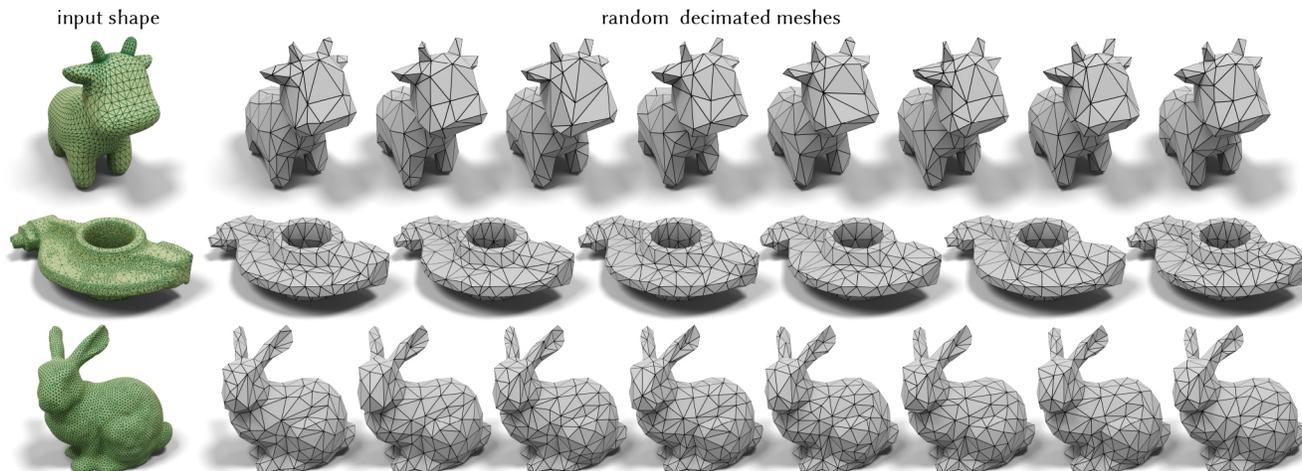


Fig. 31. We perform QSLIM with a random sequence of edge collapses to create different coarse discretizations (gray) from a single ground truth mesh (green).

says that if an edge  $e_{ij}$  connecting vertices  $i, j$  is valid, the intersection between the vertex 1-ring of  $i$  and the vertex 1-ring of  $j$  must contain only two vertices, and the two vertices cannot be an edge.

*Skinny triangles.* To prevent badly shaped triangles from causing numerical issues, we need to keep track of the triangle quality for each edge collapse. The quality of a triangle is measured by

$$Q_{ijk} = \frac{4\sqrt{3} A_{ijk}}{l_{ij}^2 + l_{jk}^2 + l_{ki}^2} \quad (3)$$

where  $A$  is the area of the triangle and  $l$  are the lengths of triangle edges. When  $Q \rightarrow 1$ , it approaches an equilateral triangle; when  $Q \rightarrow 0$ , it approaches a skinny degenerated one. For each edge, we check  $Q$  for all the neighboring faces in both UV and Euclidean domains after the collapse. By default, a valid edge requires  $Q > 0.2$  for all neighboring triangles.

#### D COMPARISON TO [Lee et al. 1998]

One possible solution to construct a bijective map between the input and the decimated model is via MAPS [Lee et al. 1998]. However, MAPS constructs the parameterization via successively removing the maximum vertex independent sets. The main reason for removing the maximum independent set is to bound the number of levels of the mesh hierarchy, but it leads to limitations such as sensitivity to the input triangulation.

One experiment to verify this is to apply subdivision remeshing presented in Sec. 4.1 in [Lee et al. 1998]. In Fig. 32 we create a stress test using a very uneven triangulation, and MAPS suffers from creating non-uniform parameterization. In contrast our successive self-parameterization enjoys the benefits of area-weighted QSLIM to obtain a more uniform parameterization.

#### E DATA GENERATION FROM RANDOM COLLAPSES

The training data for neural subdivision is a sequence of subdivided meshes where the vertex positions are computed using successive

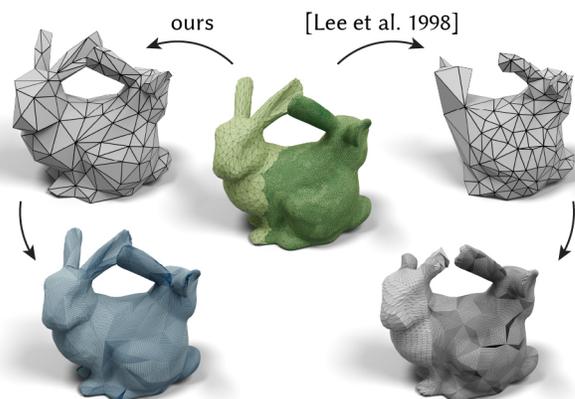


Fig. 32. We decimate the mesh down to the same number of vertices and compare our method with MAPS on the task of subdivision remeshing. Our method creates a more uniform parameterization (left), but MAPS is more sensitive to the input triangulation (right).

self-parameterization (Fig. 8). For each dense input mesh, we perform semi-random edge collapses in order to generate many different coarse meshes. The goal is to help the network to be robust to different discretizations. In Fig. 31 we show input meshes (left) can be decimated differently to get many coarse meshes that have different number of vertices and with different triangulations.

Our semi-random edge collapse starts by randomly selecting 100 edges and finding the one with the minimum *quadratic error* [Garland and Heckbert 1997] to collapse. For each edge collapse, we insert the new vertex the same way as QSLIM. We terminate the edge collapses when a randomly selected target number of vertices between 150 and 300 is reached.

#### F EXPERIMENTAL SETUP

Our experimental setup is consistent throughout the document. The training shape is presented in green in every figure. For each shape,

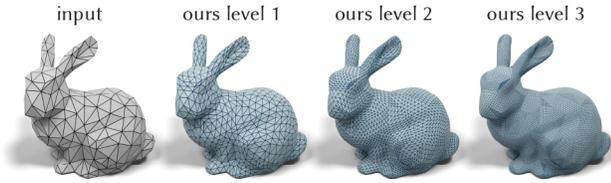


Fig. 33. Although most experiments are trained on performing 2-level subdivisions, our neural subdivision network can still be trained on more level of the subdivisions.

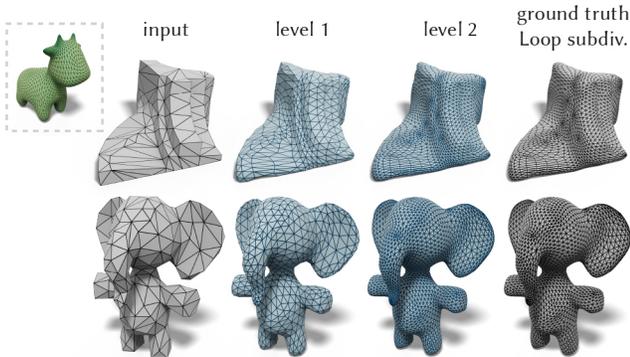


Fig. 34. When trained on meshes created by classic Loop subdivision (green), our network can reproduce the Loop scheme on new meshes, and creates visually indistinguishable results (blue) compared to the ground truth created by the classic Loop method (right).

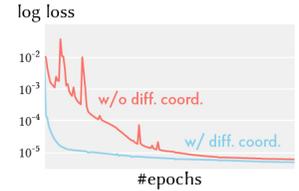
we use the parameters described in App. E to generate 200 training discretizations and train for 700 epochs. Our method can learn to produce several subdivision levels Fig. 33, but we set the number of training subdivisions to two levels for consistency across the experiments. If the experiment consists of multiple training shapes, such as the experiments in Fig. 28 and Table 2, we evenly distribute the number of training discretizations so that they still sum up to 200 discretizations in total.

## G LEARNING CLASSIC LOOP SUBDIVISION

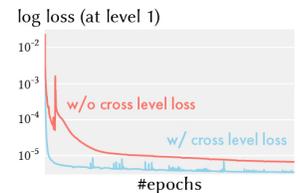
Although we have shown in Sec. 6 that neural subdivision is able to subdivide a mesh adaptively, one might be interested in seeing whether neural subdivision can also learn to reproduce classic Loop subdivision with appropriate training data. In Fig. 34, we trained our network on a sequence of meshes created with Loop subdivision. Given an original mesh, we create 200 meshes using random edge collapses, then subdivide each coarsened mesh for two levels using Loop to obtain the corresponding ground truth subdivided sequences for measuring the reconstruction loss. We see that when testing on novel meshes, the network is able to reproduce the Loop scheme to create visually indistinguishable results. The average per-vertex numerical error is just 0.3% of the bounding box diagonal.

## H ABLATION STUDIES (CONTINUED)

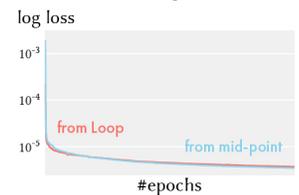
This section summarizes the ablation studies of other design decisions we made in the network design. These components are not as crucial as the components mentioned in the main text, but they still offer improvements while training. The first analysis is the influence of differential coordinates in the input (see Fig. 19). Our result in the inset indicates that adding differential coordinates can improve convergence.



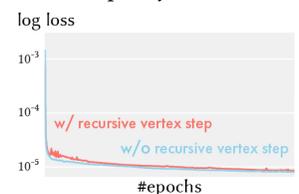
We also measure the effect of adding cross-level loss compared to only measuring the loss at the final level. In the inset, we visualize the error in the intermediate level. The result suggests that adding cross-level loss can improve subdivision results in the intermediate levels, which is important for creating meshes with different levels of detail (see Fig. 2).



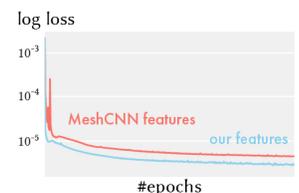
The third study is on the starting position of the predicted displacement vector as shown in Fig. 20. Specifically, we compare predicting the displacement from the mid-point of an edge with predicting displacement from the Loop-subdivided mesh. Our result in the inset suggests that using different starting positions has no influence to the quality of the output. Thus we choose the mid-point for simplicity.



The fourth study is on the number of vertex steps to perform. In Fig. 18, we can actually recursively perform the vertex step to gather information from larger rings. However our experiments in the inset indicates that recursively performing the vertex step does not offer improvements. Thus we only perform the vertex step once. We suspect that the 2-ring information on the coarse mesh (one from initialization, one from the vertex step) may already be sufficient for the network to perform subdivisions.



In MeshCNN, Hanocka et al. [2019] propose a set of features to characterize an undirected edge (via features of a flap), including the dihedral angle, two inner angles, and two edge length ratios (see Sec. 3 in [Hanocka et al. 2019]).



We tried their proposed features in our neural subdivision network. In the inset, we observe that using our features, edge vectors and the vectors of differential coordinates, converges to a better solution.