

What's the Situation with Intelligent Mesh Generation: A Survey and Perspectives

Na Lei, Zezeng Li, Zebin Xu, Ying Li, and Xianfeng Gu

Abstract—Intelligent Mesh Generation (IMG) represents a novel and promising field of research, utilizing machine learning techniques to generate meshes. Despite its relative infancy, IMG has significantly broadened the adaptability and practicality of mesh generation techniques, delivering numerous breakthroughs and unveiling potential future pathways. However, a noticeable void exists in the contemporary literature concerning comprehensive surveys of IMG methods. This paper endeavors to fill this gap by providing a systematic and thorough survey of the current IMG landscape. With a focus on 113 preliminary IMG methods, we undertake a meticulous analysis from various angles, encompassing core algorithm techniques and their application scope, agent learning objectives, data types, targeted challenges, as well as advantages and limitations. We have curated and categorized the literature, proposing three unique taxonomies based on key techniques, output mesh unit elements, and relevant input data types. This paper also underscores several promising future research directions and challenges in IMG. To augment reader accessibility, a dedicated IMG project page is available at https://github.com/xzb030/IMG_Survey.

Index Terms—Mesh generation, polygonal mesh, deep learning, neural network, survey, review

1 INTRODUCTION

1.1 Motivation

In this paper, we embark on a systematic review of publications centered on Intelligent Mesh Generation (IMG). Recognized as one of the six fundamental research directions in NASA's Vision 2030 [1], mesh generation holds a pivotal role in computational geometry and is integral to numerical simulations. IMG amalgamates machine learning with mesh generation, a blend whose significance is gaining recognition as a growing corpus of research that explores the use of neural networks to generate high-quality meshes across diverse application scenarios. Driven by various meshing objectives, IMG research has seen a surge in interest. Notably, recent years have seen the advent of many innovative algorithms. Despite the substantial research on IMG, there is a conspicuous absence of comprehensive reviews that adopt a standardized methodology to ensure significance, completeness, and impartiality. To advance research in the IMG field, an urgent need exists for a systematic overview of the current state-of-the-art IMG methods. Such an overview will aid in distilling the essential commonalities in these works, identifying current research trends, and pinpointing promising future directions.

- This research was supported by the National Key R&D Program of China (2021YFA1003003) and the National Natural Science Foundation of China under Grants 61936002 and T2225012.
- N. Lei is with the International Information and Software Institute, Dalian University of Technology, Dalian, 116620, China (Email: nalei@dlut.edu.cn).
- Z. Li and Z. Xu are with the School of Software, Dalian University of Technology, Dalian, 116620, China (Email: zezeng.lee@gmail.com, xzb0516@mail.dlut.edu.cn).
- Y. Li is with the College of Computer Science and Technology, Jilin University, Changchun, 130015, China (Email: liying@jlu.edu.cn).
- X. Gu is with the Department of Computer Science and Applied Mathematics, State University of New York at Stony Brook, Stony Brook, NY 11794-2424, USA (Email: gu@cs.stonybrook.edu).
- Corresponding author: Ying Li.

Herein, our aim is to present a comprehensive and systematic survey, complete with a detailed taxonomy and content oriented evaluation. We undertake a multi-faceted analysis, examining key algorithmic techniques and their application scope, agent learning objectives, data types, targeted challenges, along with advantages and limitations. As depicted in Fig. 1, we offer varied taxonomies for existing IMG methods, conceptualized from three perspectives: key techniques, output mesh unit elements, and relevant data types. Considering key techniques, we categorize IMG methods into deformation-based, classification-based, isosurface-based, Delaunay triangulation-based, parametrization based, and advancing front-based mesh generation. From the standpoint of output mesh unit elements, IMG methods are classified into triangular mesh, quadrilateral mesh, hybrid polygon mesh, and tetrahedral mesh generation. Lastly, we classify application input data types into point cloud-based, image-based, voxel-based, mesh-based, boundary or sketch-based, and latent variable-based mesh generation. Additionally, we present a concise overview of IMG's related areas, such as conventional neural networks for mesh learning, mesh quality metrics, and datasets. To summarize, we researched and reviewed 190 papers, which includes 113 papers focused on IMG and 77 papers on associated topics. Our paper makes four primary contributions:

- We have undertaken a comprehensive review of 113 seminal IMG papers, evaluating them based on several key aspects: underlying techniques, the scope of application, agent learning objectives, data types, targeted challenges, and advantages and limitations.
- We have categorized existing IMG methods from three distinct vantage points: key techniques, output mesh unit elements, and suitable input data types;
- We provide a broad overview of areas closely related to IMG, including conventional neural networks for

mesh learning, mesh quality assessments, and datasets;

- We encapsulate the present challenges confronting IMG and outline potential promising avenues that researchers could explore to address these challenges in their future endeavors.

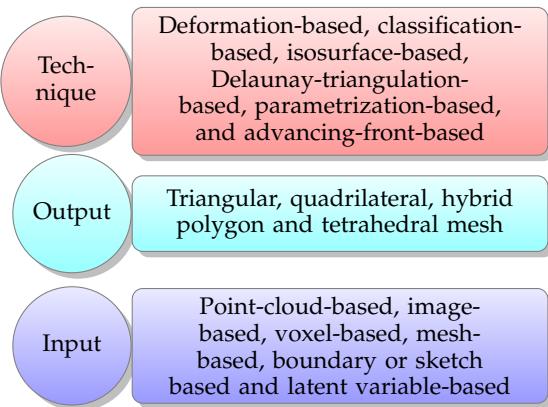


Fig. 1: Taxonomies for existing IMG methods from three views.

1.2 Related Surveys

A systematic literature review (SLR) [2], as utilized in software engineering, is conducted by using a predefined methodical series of steps. To the best of our knowledge, no SLR has been conducted in the IMG domain. In this section, we describe some related review articles [3]–[7].

Berger *et al.* [3] surveyed various point cloud-based surface reconstruction algorithms developed from 1992 to 2015. They classified the algorithms based on the type of priors that are employed, the ability to handle point cloud artifacts, input requirements, shape class, and reconstruction output. This review focuses on point cloud surface reconstruction whose output does not necessarily mesh. Note that mesh reconstruction and surface reconstruction are two different concepts. Mesh reconstruction refers to not only surface mesh but also volume mesh. However, surface reconstruction is not necessarily meshing but may also be voxel, RGBD image, signed distance field, etc. There have been other surveys covering only a specific surface reconstruction domain, such as image-based, 3D object reconstruction [4], surface remeshing [5], and advancing front-based methods [7]. All the surveys mentioned above focus on specific surface reconstruction domains but are not specialized for IMG. In addition, Xiao [6] *et al.* summarize geometric deep learning from a representational perspective. Thus, our paper is the first review of existing IMG methods.

2 APPROACH

The aim of this paper is to offer a comprehensive and systematic survey by employing a dual-pronged strategy. In the initial phase, we delve into the existing literature, sieving out IMG-relevant work using both an open-ended interpretive approach and a keyword-focused literature review. The former approach is grounded on the authors' expertise, leading to a broad and exploratory collection of works,

though it may introduce selection bias. In contrast, the latter targets specific IMG publications, offering a clear boundary, but its effectiveness relies heavily on the chosen keywords. In an effort to amalgamate the strengths of these two methods, we have subsequently applied both the keyword-centric and the open-ended interpretive approaches. In the subsequent phase, we compile and evaluate all articles gathered from the first phase based on the content extracted.

2.1 Search Strategy

The search strategy is paramount in assembling pertinent literature on a specific subject. The choice of the search string and the libraries are significant factors. We crafted a search string and selected five distinct libraries¹ to scout for relevant papers. Adhering to the SLR guidelines [2], we constructed a search string by amalgamating the keywords and their synonyms using Boolean operators.

Keywords and Boolean Operators: [(“computer graphics” OR “computational geometry”) AND (“mesh reconstruction” OR “mesh generation” OR “Surface reconstruction” OR “3D object reconstruction” OR “triangulation” OR “quadrilateral” OR “tetrahedral” OR “hexahedral”) AND (“intelligent” OR “learning” OR “data-based” OR “neural network” OR “ANN”)] OR (“mesh evaluation” OR “3D object dataset” OR “mesh neural network” OR “mesh metric”).

2.2 Content Extraction and Presentation of Findings

The existing literature offers a plethora of IMG techniques. We chose a total of 190 papers for a detailed analysis; 113 papers focus on IMG, while 77 papers delve into IMG’s related areas, encompassing classical neural networks for mesh, mesh quality metrics, and mesh datasets, etc. In light of our research questions, we extracted the following details from each selected paper: 1)The main challenges faced and significant breakthroughs achieved; 2)The fundamental concept, scope of applicability, along with pros and cons; 3)The type of input data, output mesh, and mesh quality. Furthermore, to appraise these IMG methods more objectively, we gathered the following additional information: 1)The frequency of citation for the article (number of citations per month as per Google Scholar citation); 2)The practicality of the algorithm proposed in the paper.

From the perspective of key techniques, output mesh types, and input data types, we first classified existing IMG papers and presented them in Table 1. This table also encapsulates whether the IMG is end-to-end, its principal network structure, and the learning goals of the neural network model. Table 1 provides a clear categorization of the 113 IMG papers from three unique viewpoints, allowing an intuitive understanding of how intelligence is incorporated in each IMG method. Secondly, we delved into the detailed content of each IMG paper and organized Table 2 in chronological order. This table lists the targeted challenges, advantages, limitations, and average monthly citations for each study. We also showcased some representative methods in Fig. 2, collated existing IMG papers into a line chart

1. 1) IEEE Xplore <http://ieeexplore.ieee.org>; 2) ACM <https://dl.acm.org>; 3) ScienceDirect <http://www.sciencedirect.com>; 4) SpringerLink <https://link.springer.com>; 5) Wiley <https://onlinelibrary.wiley.com/>.

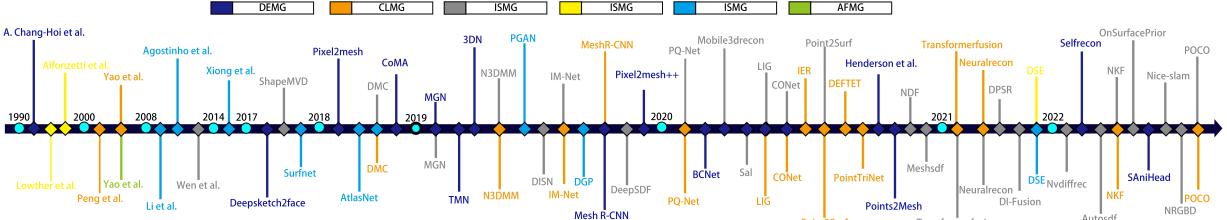


Fig. 2: Chronological overview of representative IMG methods

according to the number of papers published per year, and displayed this in Fig. 3. Given that our collection concluded in the first half of 2022, the number of IMG papers for the year has seen a notable decrease. However, Fig. 3 illustrates that IMG research is currently at its peak. Furthermore, we displayed the number of papers published per year for each technology as a stacked bar chart. For papers belonging to multiple categories, we counted them separately; hence, the sum of bar values will exceed the total number of papers.

methods do not extend beyond point clouds, images, voxels, boundaries or sketches, meshes, and latent variables.

3 CLASSIFICATION BASED ON TECHNIQUE

In this section, we categorize existing **IMG** methods into six groups, namely deformation-based, classification-based, isosurface-based, Delaunay triangulation-based, parametrization based, and advancing front-based mesh generation. This classification is grounded on the key techniques and guiding principles inherent in each method.

In this classification, we illustrate how predominant IMG methodologies adeptly amalgamate traditional mesh generation techniques with novel deep learning modules, highlighting the merits and drawbacks of each approach. We anticipate this taxonomy will help readers grasp the function of neural networks in mesh generation more intuitively. In deformation-based IMG algorithms, neural network modules primarily predict vertex locations. For classification-based IMGs, neural networks primarily function as classifiers determining the connection or occupancy of vertices. Isosurface-based IMGs employ neural networks for implicit function fitting and zero isosurface extraction. In Delaunay triangulation-based IMGs, neural networks are variably employed to predict vertex coordinates or weights. For parameterization-based IMG algorithms, neural networks learn parameterized mapping or process parameterized data. In advancing front-based IMG algorithms, neural networks are utilized to predict the forward direction of the wave and the connection type of the new node. Further elaborations are provided in the corresponding subsections below.

3.1 Deformation-based Mesh Generation

A mesh consists of polygons and is essentially a discrete representation of a continuous surface. The combinatorial nature of the polygons prevents taking derivatives over the space of possible meshing of any given surface. Thus, it is difficult for mesh processing and optimization techniques to take advantage of modular gradient descent components of modern optimization frameworks. To circumvent this problem, deformation-based IMG has attracted more attention. As shown in Fig.4, a distinguishing feature of this technology is the need for an initial mesh such as spherical or ellipsoidal template mesh. Deformation-based IMG is applicable to different data types of inputs. A target mesh based on geometric information is usually generated from images [8]–[34], point clouds [35]–[39] or voxels [40], [41].

Due to the presence of an initial mesh, this type of method reduces the difficulty of mesh generation to some extent. The neural network only needs to predict the position of the vertex because the connection relationship already exists.

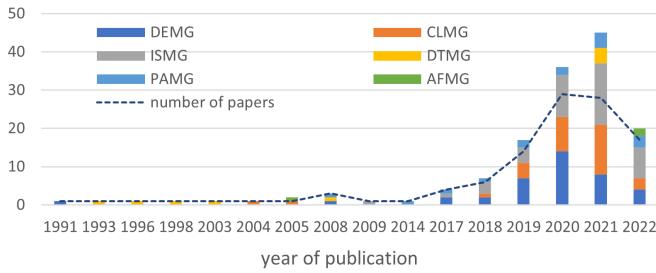


Fig. 3: Annual distribution of articles of 6 technical types.

2.3 Fundamental Concepts

In general, meshes are divided into **surface meshes** and **volume meshes**. A surface mesh is a set of polygonal faces, and the goal is to form an approximation of an object's surface. A polygonal surface mesh has three different combined elements: *vertices*, *edges*, and *faces*. This mesh can also be considered the combination of *geometry* and *topology*, where the *geometry* provides the positions of all its vertices, and the *topology* provides the information between different adjacent vertices. Mathematically, a polygonal surface mesh \mathcal{M} that contains V vertices and F faces is formed as

$$\begin{aligned} \mathcal{M} = & \{(\mathcal{V}, \mathcal{F}) | \mathcal{V} = \{v_i\}_{i=1,2,\dots,V}, \\ & \mathcal{F} = \{f_i\}_{i=1,2,\dots,F}, f_i \in \mathcal{V}^d\}, \end{aligned} \quad (1)$$

where \mathcal{V} is the vertex set, \mathcal{F} is the face set, and d denotes the number of vertices of each facet. Similarly, a volume mesh \mathcal{M} is formed as

$$\mathcal{M} = \left\{ (\mathcal{V}, \mathcal{F}, \mathcal{T}) \mid \mathcal{T} = \{t_i\}_{i=1,2,\dots,T}, \ t_i \in \mathcal{F}^k \right\}, \quad (2)$$

where k represents the number of faces of each voxel t_i . To avoid ambiguity, the definition of mesh generation

To avoid ambiguity, the definition of mesh generation is given as follows:

Definition 1. Mesh Generation: *Mesh generation is such a task: given the input data X , the mapping $g : X \rightarrow \mathcal{M}$ acts on X and outputs the surface mesh $\mathcal{M} = \{\mathcal{V}, \mathcal{F}\}$ or volume mesh $\mathcal{M} = \{\mathcal{V}, \mathcal{F}, \mathcal{T}\}$.*

In essence, the goal of mesh generation is to restore the missing geometry information or topology information in X . Through our survey, the input data types of existing IMG

The oldest deformation-based IMG work dates to 1991, and Ahn *et al.* [42] proposed a self-organizing neural network to automatically generate a nonuniform density mesh. This is a pioneering work, albeit only plane meshes with simple

TABLE 1: Taxonomy of the various types of IMG employed in the literature. DEMG, CLMG, ISMG, DTMG, PAMG and AFMG represent deformation-based, classification-based, isosurface-based, Delaunay-triangulation-based, parametrization-based, and advancing-front-based mesh generation, respectively. PC, IM, VO, ME, BS, and LA denote point-cloud-based, image-based, voxel-based, mesh-based, boundary or sketch-based, and latent variable-based IMG, respectively. Tri, Qua, Hyb and Tet represent triangular, quadrilateral, hybrid polygon and tetrahedral meshes, respectively. E2E means end-to-end, ✓ indicates satisfaction.

Article	Technique						Input						Output						Attributes		
	DEMG	CLMG	ISMG	DTMG	PAMG	AFMG	PC	IM	VO	ME	BS	LA	Tri	Qua	Hyb	Tet	E2E	Network Structure	Learning Goals		
Self-organizing [42]	✓									✓			✓	✓			✓	Self-organizing network	Vertices location		
Lowther <i>et al.</i> [43]				✓						✓			✓	✓			✓	Feed forward net	Element size		
Alfonzetti <i>et al.</i> [44]		✓								✓			✓	✓			✓	Let-It-Grow ANN	Node positions		
Alfonzetti <i>et al.</i> [45]		✓								✓			✓	✓			✓	ANN	Node positions		
Alfonzetti <i>et al.</i> [46]		✓								✓			✓	✓			✓	ANN	Node positions		
Peng <i>et al.</i> [47]	✓						✓			✓			✓	✓			✓	MLP	Coordinate z; point occupancy		
Yao <i>et al.</i> [48]	✓						✓			✓			✓	✓			✓	ANN	Node positions		
Alfonzetti <i>et al.</i> [49]				✓						✓			✓	✓			✓	Let-It-Grow ANN	Node positions		
Li <i>et al.</i> [50]							✓			✓			✓	✓			✓	RBFs network	Radial basis function		
Agostinho <i>et al.</i> [51]							✓			✓			✓	✓			✓	Self-organizing map	3D coordinates		
Wen <i>et al.</i> [52]		✓					✓			✓			✓	✓			✓	RBFs network	Radial basis function		
Xiong <i>et al.</i> [53]							✓			✓			✓	✓			✓	Sparse dictionary	Vertex position and triangulation		
DeepGarment [8]	✓									✓			✓	✓			✓	SqueezeNet	Vertex positions		
DeepSketch2Face [9]	✓									✓			✓	✓			✓	AlexNet	Vertex positions		
ShapeMVD [54]				✓						✓			✓	✓			✓	U-net	Depth and normal maps		
Surfnet [55]										✓			✓	✓			✓	Res-UNet	Geometry images		
Pixel2mesh [10]	✓									✓			✓	✓			✓	GCN/VGG-16;MLP	Vertex's coordinate/extraction characteristics		
AtlasNet [56]							✓			✓			✓	✓			✓	PointNet/ResNet+MLPs	Point position		
Li <i>et al.</i> [57]		✓					✓			✓			✓	✓			✓	U-Net	Normal, depth maps		
DMC [58]		✓	✓				✓			✓			✓	✓			✓	DMC	Occupancy and vertex displacement		
CoMA [59]	✓									✓			✓	✓			✓	Autoencoder	Point position		
3D-CFCN [60]										✓			✓	✓			✓	OctNet-based U-Net	Truncated signed distance field		
MGN [11]	✓									✓			✓	✓			✓	MLP	Vertices position		
3DN [35]	✓									✓			✓	✓			✓	PointNet/VGG	Vertices position		
TMN [12]	✓									✓			✓	✓			✓	ResNet/MLP	Vertices position and errors		
ONet [61]										✓			✓	✓			✓	ResNet/PointNet	Grid occupancy		
N3DMM [62]	✓		✓	✓						✓			✓	✓			✓	Spiral-Conv GAN	Vertices position		
PGAN [63]										✓			✓	✓			✓	WGAN	Geometry image		
HumanMeshNet [13]	✓									✓			✓	✓			✓	Resnet-18	Vertices position		
DNN [64]			✓							✓			✓	✓			✓	VGG-16	Signed distance field		
IM-Net [65]		✓	✓							✓			✓	✓			✓	IDNet	Signed distance field		
Scan2Mesh [66]		✓								✓			✓	✓			✓	3D-Conv CNN	Mesh face		
DGP [67]										✓			✓	✓			✓	MLP	Local parametrization		
Mesh R-CNN [14]	✓	✓								✓			✓	✓			✓	Mesh R-CNN	Occupancy and point position		
DeepSDF [68]				✓						✓			✓	✓			✓	MLP	Signed distance field		
Pixel2mesh++ [15]	✓									✓			✓	✓			✓	VGG; GCN	Vertex position		
PQ-Net [69]		✓	✓							✓			✓	✓			✓	Seq2Seq Autoencoder	Signed distance field		
BCNNet [16]										✓			✓	✓			✓	ResNet/GAT; Spiral-Conv	SMPL parameters; vertices position		
PolyGen [70]										✓			✓	✓			✓	Transformer-based	Predict vertices and faces sequentially		
DGTS [71]	✓									✓			✓	✓			✓	GCN	Displacement vector per face		
Neural Subdivision [72]	✓									✓			✓	✓			✓	MLP	Predict vertex position		
Mobile3drecon [73]										✓			✓	✓			✓	Res-UNet	Depth map		
Sal [74]										✓			✓	✓			✓	MLP	Unsigned distance field		
Pixel2mesh2 [17]	✓									✓			✓	✓			✓	GCN+G-Resnet	Vertex position		
Voxel2mesh [40]										✓			✓	✓			✓	GCN; 3D	Vertex position		
X-ray2shape [18]										✓			✓	✓			✓	CNN; GCN	Vertex position		
Li <i>et al.</i> [19]		✓								✓			✓	✓			✓	Resnet; 3D-GCN	Vertex position		
Meshlet [75]			✓							✓			✓	✓			✓	FC-based AE	Vertex positions		
LIG [76]		✓	✓							✓			✓	✓			✓	AE; 3D-CNN; Residual blocks	Truncated signed distance field		
CONet [77]		✓	✓							✓			✓	✓			✓	U-Net	Grid occupancy probability		
ILSM [78]		✓	✓							✓			✓	✓			✓	CGAN	Occupancy; velocity field		
IER [79]										✓			✓	✓			✓	SpareConv; MLP	Classifying triangles		
Yang <i>et al.</i> [20]	✓									✓			✓	✓			✓	VGG16; GCN; Graph attention	Vertex position		
MeshingNet [21]										✓			✓	✓			✓	FCN/Resnet	Vertex position		
Point2Surf [80]		✓	✓							✓			✓	✓			✓	MLP	Signed distance field		
DEFITET [81]										✓			✓	✓			✓	MLP;	Predict occupancy		
BTM [36]	✓									✓			✓	✓			✓	Autoencoder	Initial mesh		
PointTriNet [82]										✓			✓	✓			✓	MLP	Triangle label		
Surface Hof [22]	✓									✓			✓	✓			✓	AE	Vertex position		
REIN [83]		✓								✓			✓	✓			✓	GraphRNN	Edge prediction		
Henderson <i>et al.</i> [23]	✓									✓			✓	✓			✓	CNN	Vertex position		
SSRNet [84]										✓			✓	✓			✓	UNet; tangent convolution	Octree vertex label		
MeshVAE [85]	✓									✓			✓	✓			✓	VAE; GCN	Vertex position		
Point2Mesh [37]										✓			✓	✓			✓	GCN	Vertex position		
NMF [38]	✓									✓			✓	✓			✓	Conditional Flow/PointNet	Vertex position		
NDF [86]										✓			✓	✓			✓	IF-Nets	Unsigned distance field		
Meshsdf [87]										✓			✓	✓			✓	SplineCNN	Vertices position		
Smirnov <i>et al.</i> [26]	✓									✓			✓	✓			✓	ResNet-18	Local parameterization		
Lasr [24]										✓			✓	✓			✓	Flow/Mask Nets	Vertices position		
Transformerision [88]		✓	✓							✓			✓	✓			✓	Transformer; 3D-CNN	Learning occupancy field		
CSPNet [89]		✓	✓							✓			✓	✓			✓	CONet; PointNet	Unsigned distance field		
DASM [25]	✓									✓			✓	✓			✓	GCN; Mesh R-CNN	Vertices position		
Neuralrecon [90]		✓	✓							✓			✓	✓			✓	GRU; MLP	Truncated signed distance field		
Sa-cononet [91]		✓	✓							✓			✓	✓			✓	PointNet; CNN	Truncated signed distance field		
DEF-Met [92]										✓			✓	✓			✓	CGAN	Vertex position; signed distance field		
Vis2mesh [93]		✓								✓			✓	✓			✓	U-Net; PCconv	Visibility map		
IMLSNet [94]		✓								✓			✓	✓			✓	Octree based CNNs	Signed distance field		
Retrievalseval [95]										✓			✓	✓			✓	U-Net; Attention	Truncated distance field		
DeepIT [96]		✓								✓			✓	✓			✓	GCN	Inside/outside classification		
Iso-pool [97]										✓			✓	✓			✓	IDR	Iso-surface		
DTF [98]										✓			✓	✓			✓	MLP	Parameterization		
SAP [99]										✓			✓	✓			✓	PointCNN	Grid indicator function		
Bertiche <i>et al.</i> [27]	✓									✓			✓	✓			✓	PointCNN	SMPL parameters; vertices position		
Needrop [100]		✓	✓							✓			✓	✓			✓	OneNet	Inside/outside classification		
LMR [28]										✓			✓	✓			✓	RNN;	SMPL parameters		
NRSM [29]										✓			✓	✓			✓	Unet	Location of points		
AnalyticMesh [101]		✓	✓							✓			✓	✓			✓	VQ-VAE; transformer	Isosurface		
DHSP [102]	✓									✓			✓	✓			✓	MeshCNN; 2DCNN	Vertices position; texture map		
Neural-Pull [39]										✓			✓	✓			✓	MLP	Signed distance field		
DI-Fusion [103]										✓			✓	✓			✓	MLP based AE	Truncated signed distance field		
DSF [104]										✓			✓	✓			✓	FoldingNet	Parameterization		
LDFQ [105]										✓			✓	✓			✓	Pointnet;SpiralNet;	Direction fields		
DGN [106]		✓	✓							✓			✓	✓			✓	GNN+MLP	Tetrahedron inside/outside scores		
Hu <i>et al.</i> [30]	✓									✓			✓	✓			✓	GAN	Mesh texture		
NMC [107																					

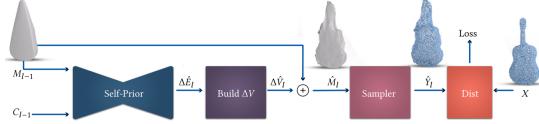


Fig. 4: Point2Mesh [37]: a deformation-based IMG method

geometry. In recent years, a sheer volume of practical methods has emerged. Most of the early deformation-based IMG works involved the deformation of the basic model to obtain the target mesh, such as clothing [8], face [9], and body [13]. These methods only generate deformation meshes of specific target objects. Later, a series of algorithms represented by Pixel2Mesh [10] introduced templates as the basic mesh, which greatly improved the practicality and generalization. The deformation-based approach is tightly integrated with GCN and MLP for better prediction of vertex positions.

However, all deformation-based models have a common drawback: they only generate meshes with the same topology as the initial mesh. Therefore, another research direction is how to adapt it to a variety of topologies. Rios *et al.* [36] proposed a very direct solution by providing multiple basic template meshes for the algorithm to automatically select. Notably, this method does not fundamentally solve the above problems but only alleviates them. To proceed, Charrada *et al.* [34] introduced a face-pruning mechanism, which iteratively adapted the topology of the mesh, using face-pruning operations, while keeping the main properties of the template, i.e., the appealing visual aspect and uniform mesh connectivity. Admittedly, the face-pruning mechanism can make the method adaptable to mesh generation for more complex topologies.

3.2 Classification-based Mesh Generation

Inspired and encouraged by the classification model in the structured data domain, many researchers have designed various effective IMG algorithms by combining classification model-based neural networks with mesh generation. These algorithms are divided into two categories: occupancy prediction [14], [32], [47], [58], [61], [65], [69], [76]–[78], [80], [81], [88]–[91], [93], [94], [96], [100], [101], [106], [107], [110], [117], [121] and mesh basic element prediction [66], [79], [82], [83]. The former voxelized the three-dimensional space and decides whether the voxel belongs to an object, which is a binary classification problem. Alternatively, occupancy prediction can be further regarded as the three classification problems of the interior, surface and exterior of the object. The latter regards mesh generation as a combination of basic elements, so the generation model only needs to judge whether the basic elements to be detected should exist, which is also a binary classification problem in essence. Examples include the classification of facet existence and local connectivity. Besides binary classification, multiclassification can also be designed to generate meshes, such as IER [79]. Fig. 5 and Fig. 6 show the IMG methods based on occupancy and basic element classification, respectively.

IMG algorithms based on occupancy prediction are represented by occupancy networks [61], which are mainly used to address the task of reconstructing mesh by giving single or multiple images. Early works locate the object surface by judging the occupancy of each grid and then reconstruct the

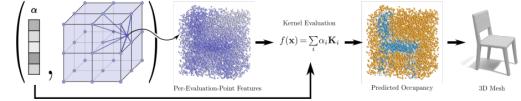


Fig. 5: NKF [110]: an occupancy classification-based IMG.

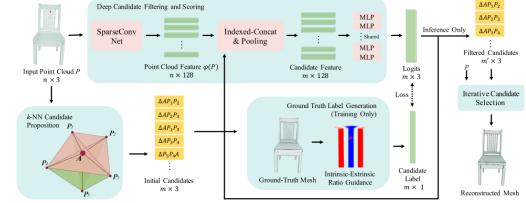


Fig. 6: IER [79]: a facet existence classification-based IMG.

mesh by marching cubes [122]. In the high-resolution case, the floating point operation proportional to the cube of the voxel resolution is unbearable. The grid occupied by the object often consists of only a small number of cubes gathered in the same area, which means that many calculations are wasted. To overcome the problem of high computational complexity, IM-NET [65] proposed a more efficient sampling approach that samples more points near shape surfaces and disregards most points far away. When addressing the task of generating meshes from point clouds, subsequent works [80], [121] focused on judging the occupancy of the point cloud instead of the whole 3D space, which directly reduces the floating-point computation.

Compared with the former, these IMG methods based on mesh basic element prediction generate meshes end-to-end. Yao [48] proposed the first algorithm to generate a mesh using a classification model, which is employed to determine whether to add new vertices and in what way. IER [79] proposed a set of candidate triangle faces by constructing a k-nearest neighbor graph on the input point cloud, where the neural network was utilized to filter out the incorrect candidates. The flowchart of IER is provided in Fig. 6. Scan2Mesh [66] and REIN [83] generated meshes by classifying the existence of edges. PointTriNet [82] iteratively applied two neural networks: a classification network predicts whether a candidate triangle should appear in the mesh, while a proposal network suggested additional candidates. The proposal network of PointTriNet avoids the need to define candidate faces in advance and removes the assumption that all points must be derived from triangular mesh vertices. Although the application of PointTriNet has expanded, the generated mesh still has many problems, such as nonmanifold and many holes.

3.3 Isosurface-based Mesh Generation

It is assumed that each point $X \in \mathbb{R}^3$ has an attribute value and that $H : X \rightarrow \mathbb{R}$ denotes implicit functions in space X . Then, the surface composed of continuous spaces with equal attribute values is referred to as an isosurface, which is defined as:

$$\text{Isosurface}(h) = \{X | H(X) = h\} \quad (3)$$

Existing works on isosurface-based IMG leverage neural networks to fit implicit functions and extract $\text{Isosurface}(0)$.

Triangle meshes are usually reconstructed by marching cubes [222], marching tetrahedra [123], Poisson surface reconstruction [124] algorithms or neural networks [58], [92], [101], [107], [117]. According to the different selected implicit functions, these IMG methods are divided into four categories: radial basis functions (RBFs) [52], occupancy fields [32], [54], [57], [58], [61], [73], [78], [84], [88], [91], [110], [119], [121], signed distance functions (SDFs) [39], [60], [64], [65], [69], [76], [77], [80], [87], [90], [92], [94]–[97], [99], [101], [103], [107]–[109], [113], [117], [120], and unsigned distance function (UDFs) [74], [86], [89], [100]. The different types of implicit functions are shown in Fig. 7.

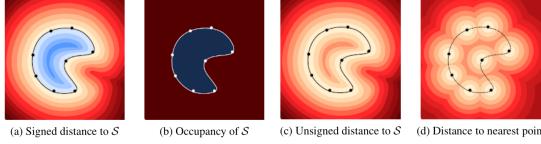


Fig. 7: Different types of implicit functions [100].

Isosurface-based IMG has the benefit of representing complex geometry and topology and is not limited to a predefined resolution. Therefore, isosurface-based IMG is favored by researchers. In 2009, Wen *et al.* [52] utilized the least squares radial basis function-based neural network to estimate the coefficient on each surface sample and constructed a continuous implicit function to represent a 3D surface. The network also overcame the problem of numerical ill-conditioning and overfitting of traditional RBF reconstruction and offered a tool to utilize fewer samples to reconstruct models and make geometry processing feasible and practical. Afterward, Lun *et al.* [54] used a decoder network to fit the foreground probability function with the occupancy fields; it is defined as follows:

$$O : X \rightarrow \{0, 1\}, X \in \mathbb{R}^3. \quad (4)$$

Recently, many works based on occupancy prediction emerged. Liao *et al.* [58] proposed differentiable Marching Cubes by predicting the probability of occupancy for each voxel. To improve voxel occupancy prediction's calculation and memory efficiency, Mescheder *et al.* proposed occupancy networks [61]. Currently, various occupancy-based methods with different network architectures and training strategies have been proposed to continuously improve the efficiency, robustness, and accuracy [32], [57], [73], [78], [84], [88], [91], [110], [119], [121]. The other two commonly utilized implicit functions are SDF and UDF. Cao *et al.* [60] introduced a cascaded, 3D, convolutional network architecture that learned SDF from noisy and incomplete depth maps in a progressive, coarse-to-fine manner. Wang *et al.* [64] presented a deep implicit surface network that generated a high-quality, detail-rich, 3D mesh from a 2D image by predicting the underlying SDF. Among these methods, IMG methods of utilizing Neural Radiance Fields (NeRF) show promising results in reproducing the appearance of an object or scene [120], [125], [126]. For them, NeRF is used to predicate SDF or density fields. Then the mesh can be extracted from the SDF using the marching cubes algorithm.

Although neural implicit representations gained popularity in 3D reconstruction due to their expressiveness and

flexibility, the nature of neural implicit representations produces slow inference time and requires careful initialization. To solve these problems, Peng *et al.* [99] introduced a differentiable point-to-mesh layer using a differentiable PSR [124], which bridged the explicit 3D point representation with the 3D mesh via the implicit indicator field, enabling end-to-end optimization. Its flowchart is shown in Fig. 8. For UDF, Atz-

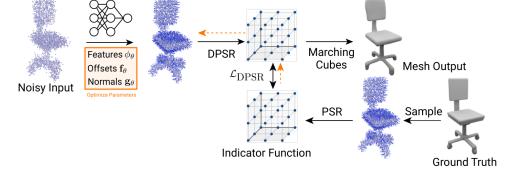


Fig. 8: SAP [99]: an isosurface-based IMG method

mon *et al.* [74] introduced a sign-agnostic learning approach for learning implicit shape representations directly from raw and unsigned geometric data. In [86], neural distance fields were proposed to predict the unsigned distance field for arbitrary 3D shapes given sparse point clouds. To estimate local geometric such as the normals and tangent planes by implicit representation, Venkatesh *et al.* [89] utilized an efficient implicit representation referred to as the closest surface point representation, which was utilized to solve the PDE-related problems on surfaces [127]–[132]. To process highly sparse input point clouds, NeeDrop [100] introduced a statistically based self-supervised approach to estimate occupancy functions directly from point clouds.

The family of isosurface approaches is sometimes limited by their sensitivity to noise, outliers, and nonuniform sampling. In addition, solving equations on a large scale in implicit methods can be time-consuming. There is a cubic or quadratic relation between resolution and run time as well as memory usage, which limits its application value.

3.4 Delaunay triangulation-based Mesh Generation

Delaunay triangulation [133] is a widely employed mesh reconstruction technology that can connect point clouds into triangular meshes. Delaunay triangulation has many excellent properties, such as maximizing the minimum angle characteristic. Regardless of which region the Delaunay triangulation is constructed from, the final generated triangular mesh is unique. Inspired by Delaunay triangulation, some researchers have tried to integrate it into IMG to generate high-quality and manifold meshes [43]–[46], [49], [98], [104]. Among these methods, early works [43]–[45],

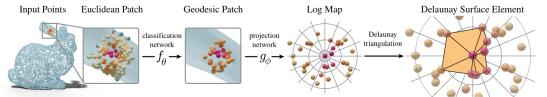


Fig. 9: DSE [104]: a Delaunay triangulation-based IMG

[49] predict the position of vertices using neural networks and then generate high-quality meshes using Delaunay triangulation [133]. These methods generate 2D planar meshes with given boundaries or coarse meshes. However, in applications, a three-dimensional surface often needs to be addressed. Alfonzetti [46] utilized the 3D Delaunay algorithm

to generate a tetrahedral mesh from a 3D point cloud. Song *et al.* [93] combined the classification model and adaptive Delaunay algorithm to realize the mesh reconstruction of large indoor and outdoor scenes. Rakotosaona *et al.* [104] projected a 3D point into the 2D parameter space by a neural network, triangulated the 2D plane point using the Delaunay algorithm, and then pulled the point cloud back to the 3D surface by inverse mapping. The algorithm pipeline is shown in Fig. 9. Furthermore, to realize end-to-end differentiable triangulation, Rakotosaona *et al.* [98] proposed a differentiable weighted Delaunay triangulation.

3.5 Parametrization-based Mesh Generation

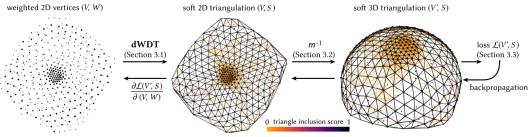


Fig. 10: DST [98]: a parametrization-based IMG method

As [134] remark, parameterization is a correspondence between a surface mesh embedded in 3D and a simple 2D domain referred to as the parameter space. Generally, parameterization is expected to be bijective, which is a classic problem in computer graphics and geometry processing with multifarious applications, such as mesh generation, texture mapping, and shape correspondence. Considering the convenience of mesh generation in parameter space, some researchers have tried to combine parameterization with machine learning modules and then invented a series of IMG methods based on parameterization. According to the different participation modes of parameterization, we broadly divide them into two categories. The **first category** is geometry image-based IMG. The authors map the surface to the parameter plane by the traditional parametric method [135] to obtain a geometry image, process the geometry image using the deep learning model, and then convert the geometry image to a mesh. The **second category** is parameterization learning-based IMG, which learns the mapping from the parameter space to the target space. Then, meshes are generated in the parameter space by the traditional or machine learning method.

For the **first category**, Sinha *et al.* [55] developed a procedure to create geometry images that represent the shape surface of a category of 3D objects. Then, the authors utilized these geometry images for category-specific surface generation by developing a variant of deep residual networks. Li *et al.* [63], [118] proposed a prediction generative adversarial network and prediction-compensation generative adversarial network to learn a joint distribution of geometry and normal images for generating meshes. They aimed to generate meshes with two generative adversarial networks (GANs) [136] and achieved good performance in face mesh generation. The advantage of this type of approach is that 3D objects can be directly processed by 2D grid convolution. However, angle and area distortion is often disregarded. Surfnet [55] reduced the effect of area distortion to some extent by making different objects of the same class have the same geometry image. For the **second category**, the main idea was to fit a mapping from the parameter space

to the object space or vice versa by a neural network. Early work [50] promoted the planar mesh to the surface by fitting the bivariate function. In [51], the mapping from a 2D triangular lattice to a surface in 3D space is learned to lift the mesh in parameter space to the target surface. Other researchers try to project the 3D point cloud onto a 2D plane, generate meshes on the plane, and then pull it back. Considering the complexity and difference in surface topology, part of the work chooses to fit local homeomorphic mapping [26], [56], [67], [75], [98], [104], [105]. The other part directly fits the global mapping by assuming the surface type [53], [102], [111], [115]. Fig. 10 shows a local parameter learning method, DST [98], which decomposed the source into local patches and then performed per-patch meshing. The method employed parameterization m , a bijective and piecewise differentiable mapping, to project 3D surfaces into 2D parametric space. Then, the 2D vertices were lifted to form a soft 3D triangulation on the surface using m^{-1} . The advantage of DST is to ensure that the generated mesh is 2-manifold, and the limitation is the boundary artifact caused by partitioning, which is a common problem for all local parameterization methods.

3.6 Advancing front-based Mesh Generation

The advancing front method is a greedy algorithm that gradually generates mesh nodes from the boundaries to the interior and recursively executes. Each recursive process is divided into three steps: first, a line segment is selected from the generation segment front set that splits the meshed domain and unmeshed domain, where the selected segment is referred to as the base segment, as it will form the basis for the creation of a new triangle element; second, a new mesh node or an existing mesh node is connected to the base segment to generate a new triangle element; last, the generation segment front set and triangle elements are updated.

Attracted by the classical advancing front method, some researchers have tried to combine it with machine learning modules to design new practical algorithms. Early advancing front method can generate simple meshes, but it involves many inefficient calculations [137], [138]. Yao *et al.* [48] combined new vertex position prediction and classification to generate quadrilateral meshes. RLQMG [116] combined the advancing front method with reinforcement learning, which generated wavefronts and new points by state representation and action representation, respectively. Lu *et al.* [114] used deep learning to learn to advance direction and step. The main limitation of this type of method is that the quality of the generated mesh cannot be guaranteed (e.g., too many singularities). In addition, the existing advancing front-based methods only generate planar meshes.

4 CLASSIFICATION BASED ON UNIT ELEMENTS

Based on the type of output mesh primitive, we categorize IMGs into triangular, quadrilateral, hybrid polygon, and tetrahedral mesh generation. This classification allows us to identify the type of mesh generated by each method, thus enabling users to select the content that best suits their needs. As demonstrated in TABLE 1, the majority of current IMG methodologies produce triangular meshes. This trend

is logical given that triangular meshes are still the prevailing format, supported by a wealth of tools and theoretical knowledge. This observation also underscores that fields other than triangular meshing, such as quadrilateral and volumetric meshing, are areas ripe for further exploration.

4.1 Triangular Mesh Generation

Triangle meshes are arguably the most predominant surface mesh representation. Its popularity is derived from its simplicity, flexibility, and the existence of many data structures for efficient mesh navigation and manipulation. With an increase in publicly available triangular mesh datasets and further development of deep learning techniques, many methods have been proposed to generate triangular meshes of given point clouds, images or other forms of data.

Notably, intelligent triangular mesh generation algorithms attracted the interest of researchers as early as the 1990s. The main focus of this period was density-adaptive triangular mesh generation algorithms [42]–[46], [49], whose main feature is the generation of triangular meshes of desired nonuniform density on an initial coarse mesh following the guidance of the density function learned by neural networks. However, the above methods mainly focus on planar meshes and rarely consider the surface mesh generation of 3D objects. Therefore, to improve the practicability of the algorithm to adapt to real-world inputs, more researchers have dedicated themselves to investigating 3D mesh generation. Peng *et al.* [47] explored the feasibility of neural networks in triangular mesh reconstruction and 3D object representation. Although this approach is very simple, it realizes lifting from 2D to 3D by predicting the z coordinate.

As 3D data acquisition equipment rapidly develops, triangle mesh generation with 3D data as input is becoming an important branch of IMG. Triangle mesh generation broadly falls into two types: (1) explicit methods, which directly recover a triangle mesh from the input points, and (2) implicit methods, which are aimed at recovering a volumetric function whose zero level-set encodes the surface. For the former, we consider the point cloud as a degenerate mesh (i.e., a mesh without the connection relationship between two points). Therefore, the selection of vertices and prediction of the correct adjacency is needed. Specifically, Scan2mesh [66] generated triangle meshes by joint prediction of point positions and connection relations. However, using the fully connected graph as the candidate set of edges makes the edge information redundant. Daroya *et al.* [83] presented REIN, an RNN-based network to generate meshes with varying numbers of vertices by sequential edge prediction. Some methods based on template deformation, as mentioned in Section 3.1, are also equivalent to initializing the point connection relations. For the latter, surface triangle meshes are often reconstructed with the marching cubes [122], marching tetrahedra [123] or Poisson surface reconstruction [124] algorithms. However, these methods cannot maintain sharp features and are not end-to-end. Therefore, some researchers have explored differentiable marching cube algorithms to maintain sharp features. Liao *et al.* [58] demonstrated a differentiable alternative to achieve end-to-end training. Liu *et al.* [94] proposed a deep moving least squares method to generate triangular

meshes based on iso-surfaces. Chen *et al.* [107] redesigned the marching cube algorithm and constructed some tessellation templates, allowing better preservation of sharp features. Chen *et al.* [117] improved the dual contouring algorithm [139] to eventually allow for end-to-end training.

Another branch is the triangular mesh generation method based on 2D images or features, which is more ambiguous than the method based on 3D data and demands more robustness prior to assist mesh generation. The reconstruction of occlusion information and the estimation of depth information become particularly important. Some deformation-based methods and parameterization-based methods have become the mainstream methods to address the abovementioned cases. For the former, the image is considered the information that guides the deformation of the template. The typical representative is the Pixel2Mesh series [10], [15], [17], [31]. For the latter, we understand 2D features as a parameterization of 3D object surfaces. The geometry image-based [55], [118] approach deserves to be noted.

4.2 Quadrilateral Mesh Generation

Due to the appealing tensor-product nature and smooth surface approximation of quad meshes, quad meshing techniques are usually preferred over triangle meshes in many applications, such as texturing, simulation with finite elements, and B-spline fitting. To take advantage of the powerful representation and generalization capabilities of deep learning, some researchers have also tried to combine deep learning with quad mesh generation and have obtained good solutions [26], [29], [42], [48], [105], [111], [115], [116].

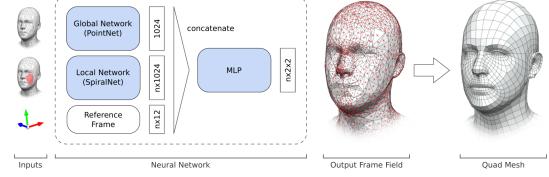


Fig. 11: LDFQ [105]: an IMG method for quadrilateral mesh.

Following the traditional method of classification, we categorize deep learning-based quadrilateral mesh generation into **direct** [29], [42], [48], [115], [116] and **indirect** [26], [105], [111] methods, according to whether intermediate model representation is needed. In these direct IMG methods, Ahn *et al.* [42] proposed a self-organizing neural network that achieves automatic nonuniform density quad mesh generation by deforming the initial mesh. Yao *et al.* [48] proposed an artificial neural network-based element extraction method for automatic finite element quad mesh generation. The authors designed reasonable node insertion types and then employed neural networks to predict the node positions and insertion types. The mesh boundary is continuously advanced by several iterations of updates. Pedone *et al.* [29] generated a database of small deforming rectangular meshes. Chen *et al.* [115] introduced a differential method for structured mesh generation in an unsupervised manner. The method takes boundary curves as input, employs a well-designed neural network to analyze the potential meshing rules, and outputs the mesh with a desired number of cells. To overcome the difficulties

of conventional methods in achieving the balance between high-quality mesh and computational complexity, Pan *et al.* [116] proposed a reinforcement learning-based method for automatic quadrilateral mesh generation. Unfortunately, all these methods can only handle simple planes or surfaces.

For indirect methods, Smirnov *et al.* [26] learned a special shape representation: a deformable parametric template composed of Coons patches. Given a raster image, first, the system infers a set of parametric surfaces that realize the input in 3D. Then, quadrilateral meshes are generated using a template. Deng *et al.* [111] presented the Sketch2PQ system, which uses stroke lines, depth samples, and the visible and occluded region masks induced from the sketch as input. First, Sketch2PQ calculates a direction field and B-spline surface as the intermediate model representation. Then, a quad mesh is extracted from the B-spline surface and CDF via geometry optimization. Another interesting work is LDFQ [105], which learns cross fields from a triangular mesh to generate a quadrilateral mesh. Their network can infer frame fields that resemble the alignment of quads. This rich guidance information is able to guarantee the generation of correct and high-quality quad meshes. The pipeline of LDFQ is depicted in Fig. 11.

4.3 Hybrid Polygon Mesh Generation

In practical applications, due to various nonideal conditions, it is often impossible to generate only a single basic facet. Therefore, some researchers have designed hybrid mesh generation algorithms out of active or passive intention. However, due to the limitation of their application, there are only a few types of IMG with hybrid meshes as the generation target. In the literature that we searched, only BP-ANN [114], AnalyticMesh [101] and PolyGen [70] belong to this category. BP-ANN generates anisotropic quadrilateral and isotropic triangular meshes by the advancing front method, which improves the level of automation and efficiency of hybrid mesh generation. AnalyticMesh marches among analytic cells to recover the exact mesh of the closed, piecewise planar surface captured by an implicit surface network. The algorithm is applicable to more advanced MLP architectures, including those with shortcut connections and max pooling operations, which support a richer set of architectural designs for learning and exactly meshing complex surface shapes. Polygen more compactly represents the surface of the object with different polygons.

4.4 Tetrahedral Mesh Generation

Tetrahedral mesh generation is an important branch of mesh generation. Unlike the surface mesh, tetrahedral mesh generation simultaneously divides the surface and the interior of the object. Tetrahedral mesh plays an irreplaceable role in numerical simulation, so it has received the attention of many researchers. However, due to the complexity of the problem and the incompleteness of the basic theory, intelligent tetrahedral mesh generation is still lacking. In the literature that we selected, the tetrahedral mesh generation process has been partially [46] or fully [81] replaced by deep learning modules. In [46], Alfonzetti *et al.* proposed a neural network generator for tetrahedral meshes. Starting from an initial moderately coarse mesh, the generator grows

the mesh up to the user-specified number of nodes by a node probability density function. DefTet [81] is optimized for both vertex placement and occupancy and is differentiable with respect to standard 3D reconstruction loss functions. DefTet offers several advantages over prior work and can

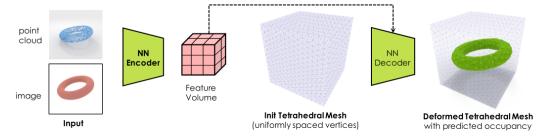


Fig. 12: DefTet [81]: an IMG method for tetrahedral mesh.

output shapes with arbitrary topology, using the occupancy of the tetrahedrons to differentiate the interior of the object from the exterior of the object. DefTet also represents local geometric details by deforming the triangular faces in these tetrahedrons to better align with the object surface, thus achieving high-fidelity reconstruction at a significantly lower memory footprint than existing volumetric approaches. DefTet accepts point clouds or images as input data to generate meshes. The algorithm flowchart of DefTet is shown in Fig. 12.

5 CLASSIFICATION BASED ON DATA TYPES

According to the type of applicable input data, we categorize IMG into categories such as point cloud-based, image-based, voxel-based, mesh-based, boundary or sketch-based, and latent variable-based mesh generation. This classification allows us to identify the type of input data that each method requires, thereby enabling users to choose the method that is most relevant to their specific data types or use-cases. As illustrated in TABLE 1, the majority of existing IMG methods generate meshes from either images or point clouds, likely due to the relative ease of obtaining these types of data compared to others. Further details are provided in the subsequent subsections.

5.1 Point Cloud-based Mesh Generation

Point clouds, as a major representation of 3D data, have been widely employed in fields such as autonomous driving and AR. Point clouds naturally have depth information and are not affected by ambient lighting conditions compared to RGB images. Reconstructing a mesh from a point cloud is a long-standing problem in computer graphics. Recently, various approaches have been developed to reconstruct shapes for an array of applications [3]. These methods are divided into two categories: direct estimation of point locations and connection relationships, as shown in Fig. 13, and implicit surface reconstruction based on isosurfaces. This division is illustrated in section 4.1. Next, we introduce these methods in two aspects: (1) the challenges of point cloud-based mesh generation and (2) the learnable priors and related methods proposed to cope with these challenges.

The challenges of point cloud-based IMG methods are derived from both data and mission objectives. For the data side, although point clouds are easily accessible 3D data, they are often sparse and noisy or even incomplete. To address these challenges, researchers have proposed a variety of approaches [39], [50], [75], [83], [97], [100], [110], [113]. Li

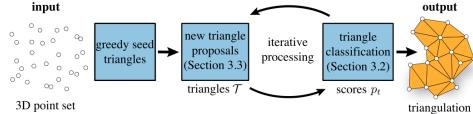


Fig. 13: PointTriNet [82]: a point cloud-based IMG method.

et al. [50] was the first to consider the problem of generating meshes from incomplete point clouds based on neural networks. Badki *et al.* [75] handled sparse or noisy point clouds by learning local geometric priors while ensuring the consistency of the global geometry. Boulch *et al.* [100] and Ma *et al.* [39], [113] reconstructed highly accurate meshes from sparse point clouds by pulling query 3D locations to their closest points on the surface. For the task objective, large scene mesh reconstruction and manifold mesh generation are the focus of this field. Great progress has been made in mesh reconstruction of large scenes [77], [91], [93], [95], [106], [110], but it is still a long way from practical application. Manifold mesh generation [98], [99], [104], [107], [117] is mainly achieved by combining Delaunay triangulation or the Marching Cubes algorithm. For example, Chen *et al.* [107] introduced the first Marching Cubes based approach capable of recovering sharp geometric features. Rakotosaona *et al.* [104] leveraged the properties of 2D Delaunay triangulations to construct a 3D mesh.

5.2 Image-based Mesh Generation

The goal of image-based mesh generation is to infer the 3D mesh of objects or scenes from one or multiple 2D images. Recovery of the lost dimensions from just 2D images is an ill-posed problem and fundamental to many applications, such as robot navigation, 3D modeling and animation, industrial control, and medical diagnosis. The rapid development of deep learning techniques, and more importantly, the increasing public datasets, accelerate the evolution of this subfield. Despite being recent, these methods have demonstrated exciting and promising results on various tasks [4]. The following is a detailed introduction.

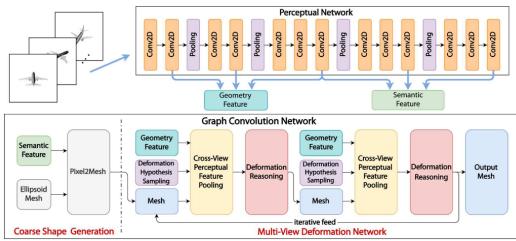


Fig. 14: Pixel2Mesh++ [15]: an images-based IMG method.

One of the immediate difficulties faced by image-based methods is how to maintain and represent the geometric and topological information of objects. Sinha *et al.* [55] constructed networks to learn geometric images with images as input to generate 3D meshes. The geometric information of the shape is preserved by parameterization learning. Wang *et al.* [10] used the deformation-based approach to restore geometry and topology. Afterward, there were many deformation-based methods [15], [17]–[20]. Wen *et al.* [15] used multiple images to build geometric details, as

shown in 14. Tong *et al.* [18] employed specific templates to provide more geometric details. Although there are many excellent works, generalization is still a weakness of these deformation-based IMG technologies.

Other researchers have turned their attention to the reconstruction of textured meshes. Although RGB images provide some texture information, the task is more challenging than purely generating a mesh. Henderson *et al.* [23] focused on the problem of generating a 3D texture mesh from 2D images. Munkberg *et al.* [108] simultaneously optimized topology, materials, and illumination from multiview image observations. Admittedly, the interpretability of IMG is enhanced when exploring the correspondence between 3D mesh textures and 2D images. In addition, mesh generation based on image sequences and videos is also worthy of attention [24], [28], [29], [33], [88], [90].

5.3 Voxel-based Mesh Generation

A voxel is a data structure that uses a fixed-size cube as the minimum unit to represent a 3D object. A voxel is a traditional method for storing volume data, providing density, opacity, normals, and other information. The index of the voxel grid provides location information. The major drawback of voxel modeling is that the storage and calculation of voxels require a large memory resource.

Similar to point clouds, a class of voxel-based methods uses occupancy and isosurfaces to generate the mesh implicitly. Mescheder *et al.* [61] used occupancy networks for 3D resolution enhancement to better reconstruct objects. Voxel2Mesh [40] utilized a network to extract features from voxelized MRI brain images and CT liver scans and relied on deformation networks to obtain a triangular mesh of the target object. Peng *et al.* [77] combined convolutional encoders with implicit occupancy decoders and achieved detailed reconstruction of objects and 3D scenes. The authors encode the input into 2D or 3D Voxel grids which are processed using convolutional networks and then decoded into occupancy probabilities via a fully connected network. DMTet [92] is a deep 3D conditional generative model that can synthesize high-resolution, 3D shapes using simple user guides such as coarse voxels. The model marries the merits of implicit and explicit 3D representations by a differentiable marching tetrahedra layer. This combination allows joint optimization of the surface geometry and topology as well as the generation of the hierarchy of subdivisions by using reconstruction and adversarial losses.

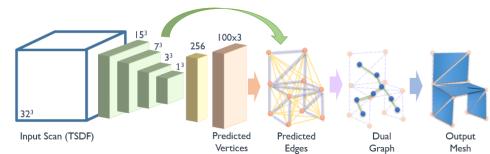


Fig. 15: Scan2Mesh [66]: a voxel-based IMG method.

As an intermediate product of iso-surface extraction, the voxelized distance field is also often employed as input to generate meshes [66], [107], [117]. Scan2Mesh [66] took the voxelized truncated signed distance field (TSDF) as input. Its framework is visualized in Fig. 15. Scan2Mesh is composed of two main components: first, a 3D-convolutional

and graph neural network to jointly predict vertex locations and edge connectivity; and second, a graph neural network to predict the final mesh face structure. NMC [107] used a signed distance field as input and was then trained to reconstruct the zero-isosurface of an implicit field while preserving geometric features such as sharp edges and smooth curves. The main limitations of NMC are that it is sensitive to rotation and cannot avoid self-intersection of meshes. NDC [117] is a data-driven approach to mesh reconstruction based on dual contouring. NDC can be trained to produce meshes from binary voxel grids, signed or unsigned distance fields, or point clouds and can produce open surfaces in cases where the input represents a sheet or partial surface.

5.4 Mesh-based Mesh Generation

In IMG, mesh-based input is mainly utilized in two tasks: mesh optimization [44]–[46], [49], [112] and quadrilateral mesh generation [105]. Early deep learning-based mesh generation tended to take a coarse mesh as input and refine it to meet the geometric and topological information of the object by neural networks. In [44], Alfonzetti proposed an automatic mesh generator that is the Let-It-Grow neural network. Starting from a rough mesh of triangles with a small number of vertices, this algorithm increases the number of vertices until a user-selected value is reached. The vertex growth is driven by a predefined probability density function. In Alfonzetti's follow-up work [45], [46], [49], he followed the idea in [44] and proposed different solutions. Liu *et al.* [72] proposed a novel framework for data-driven coarse-to-fine geometry modeling. Hahner *et al.* [112] proposed an autoencoder that handled semiregular meshes of different sizes and topologies. This algorithm can reconstruct meshes with high quality and generalize to the dynamics of unseen time sequences. Dielen *et al.* [105] proposed a neural network that inferred frame fields from unstructured triangle meshes. Afterward, a quad mesh is reconstructed from the frame field by an existing parametrization based quadrangulation method.

5.5 Boundary- or Sketch-based Mesh Generation

Compared to point clouds or images, boundaries and sketches are two types of data that are more concise and easier to obtain. Therefore, boundaries and sketches are popular input data forms for mesh generation tasks. The boundary defines the area covered by the mesh, gives the initial position and growth direction for the mesh and is a common input for advancing front-based mesh generation methods. Sketches depict rich geometric features of 3D shapes, such as silhouettes, occluding and suggestive contours, ridges and valleys and hatching lines, and thus provide a succinct and intuitive method for mesh generation. In the IMG that we collected, boundary-based mesh generation methods [21], [42], [43], [48], [114]–[116] are usually employed for 2D planar meshes, while sketch-based methods [9], [26], [41], [54], [57], [78], [111] are generally utilized for 3D object surface meshes.

For boundary-based IMG, most methods generate 2D meshes by the advancing front technique [48], [114], [116] or by solving partial differential equations with given initial boundary conditions [21], [115]. In addition, A. Chang-Hoi *et al.* [42] and D. Lowther *et al.* [43] generated meshes by

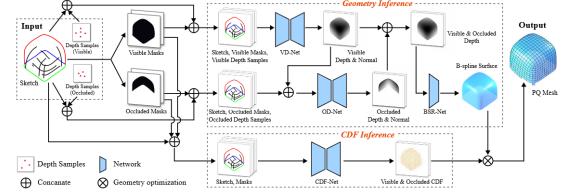


Fig. 16: Sketch2PQ [111]: a sketch-based IMG for quad mesh.

deformation and point prediction, respectively. A sketch can be considered a very sparse image. Therefore, additional information (e.g., depth information) or global and local constraints are needed. Lun *et al.* [54] divided sketch-based generation into two stages: first, generating depth and normal images from a multiview sketch and, second, generating a surface from the depth and normal images. Li *et al.* [57] used CNNs to infer the depth and normal maps that represent the surface, with an intermediate layer that models the curvature direction field and produces a confidence map to improve robustness. Deng *et al.* [111] generated quadrilateral meshes using visible and occluded boundaries, contour, and feature lines, and some depth samples. The boundary and the feature lines collectively determine the style of mesh, as shown in 16.

5.6 Latent Variable-based Mesh Generation

As the two dominant generation models, GANs and variational autoencoders (VAEs) [140] naturally exert their power in IMG. It is well known that both models can generate images directly from latent variables (random noise) after training. Inspired by the success in image generation, some researchers have conducted mesh generation research based on GANs and VAEs. Considering that only a latent variable of latent space is needed to generate mesh after model training, this paper classifies this kind of method as latent variable-based mesh generation [59], [62], [63], [69], [85], [118]. Specifically, Li *et al.* [63], [118] used the GAN model to generate geometry and normal images from a latent variable and then generated a mesh by postprocessing. Some other works [59], [62], [85] selected the VAE model to reconstruct, deform, or interpolate the mesh. Pq-Net [69] enabled the sequence generation of parts using a conditional variational autoencoder.

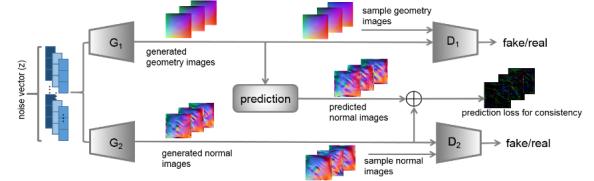


Fig. 17: PGAN [63]: a latent variable-based IMG method.

6 EVALUATION

Generally, evaluation methodologies for various IMGs fall into two major categories: qualitative and quantitative. Our qualitative analysis focuses on the content perspective of all IMG methods, including targeting challenges, advantages, limitations, and average monthly citations. On the other

hand, quantitative evaluations typically emphasize specific metrics such as time complexity and mesh quality. In this section, we undertake a thorough and comprehensive qualitative analysis based on the extracted content from all selected methods. However, due to the vastly different target tasks and datasets across various IMG methods, the metrics they employ differ significantly. Consequently, we refrain from a quantitative comparison of all IMG methods, instead providing a list of commonly used metrics.

6.1 Content evaluation

We extracted the key challenges, advantages, and limitations of the proposed algorithm from the reviewed papers and sorted them in Table 2 in chronological order. The key challenges are often the goal of a paper and embody its value to a certain extent. Advantages are often reflected in effective solutions to challenges. The limitations point out the shortcomings of the algorithm and indicate directions for future research. In specific applications, these three aspects also provide guidance for us to choose suitable algorithms.

In the summary of advantages and limitations, we did not use phrases such as higher accuracy, faster speed, or fewer memory requirements. With the development of IMG technology, the advantages claimed in the original paper are often untenable. Therefore, we try to give a more objective evaluation based on the mission objectives and application scenarios. In addition, we calculated and displayed the average monthly citations (AMC) of each paper, which to some extent reflects the impact of the corresponding work.

6.2 Metrics

In our early attempt at quantitative analysis, we experienced complicated and arduous data collation and analysis. Although we ultimately failed, we also obtained an unexpected discovery that the lack of benchmarks is a challenge hindering the development of IMG. The benchmark here includes unified test data and metrics. For quantitative analysis, most of the methods create their own test data. Even if the dataset is the same, the test data will vary due to different preprocessing. Moreover, the difference in metric selection also directly leads to the inability of quantitative comparisons. Due to the above insurmountable issues, we abandoned the attempt to make a comprehensive quantitative comparison of existing IMG methods. Next, we list and explain the commonly employed metrics in IMG so that readers can understand the challenges of comprehensive comparison in this field. Based on what metrics evaluate, we classified these metrics into error and feature preservation metrics, topological metrics, and perceptual metrics.

Error and feature preservation metrics. In IMG, the chamfer distance (CD) [141] is the most prevalent metric to measure the error between the reconstructed mesh \mathcal{M}_{gt} and the target surface \mathcal{M}_{pred} . The CD is defined as the average of the minimum distance between two point clouds. The coordinate range of points, density of points, and definition of the distance norm will directly affect the value of the CD. Whether or not the coordinates are normalized, the number of sampling points on the mesh, and whether the distance is the L_1 or L_2 norm, yield different CD values, making comparative analysis impossible. F score [10] is defined as

the harmonic mean of precision and recall, where recall is a fraction of points on \mathcal{M}_{gt} that lie within a certain distance to \mathcal{M}_{ped} , and precision is the fraction of points on \mathcal{M}_{ped} that lie within a certain distance to \mathcal{M}_{gt} . The F score mainly captures significant mistakes and disregards small mistakes that significantly contribute to visual artifacts. However, different threshold and distance selection strategies directly lead to different F score values. The normal consistency score [61] measures the consistency of the normal of the 3D model surface, that is, the normals of adjacent faces should be smooth and continuous, rather than obviously discontinuous or inconsistent. The test data must contain accurate normal vectors and involve complex projection, so the application scenario is limited.

Topological metrics. Watertightness and manifoldness are two topological metrics. Water tightness [82] denotes the percentage of edges that have exactly two incident triangles, and manifoldness [82] represents the percentage of edges with one or two incident triangles. These two metrics measure the mesh from the perspective of topology but disregard the geometry and features of the target surface.

Perception metric. As [65], [142] remarks, metrics such as the mean squared error and CD do not account for the visual quality of the object surfaces. For visual perception, low-frequency errors in shapes are less noticeable than high-frequency errors. To overcome the shortcomings of traditional quality metrics, deep learning measurement methods based on visual similarity or perceptual quality [143]–[149] have emerged. However, the generalization and applicability of these models need to be improved, so they are rarely employed in quantitative comparative analysis.

As mentioned above, we cannot use the existing results to compare all these IMG methods, nor can we compare them directly by experiments. It is very challenging to conduct a comparative analysis of these IMG methods. Nevertheless, to facilitate researchers in choosing papers of their own interest, we have provided some imperfect qualitative comparison results and comparative relationship diagrams in the supplementary materials and the project repository.

7 PERIPHERAL

7.1 Common Datasets

Data serves as the cornerstone of IMG development. The variety of inputs underscores the diversity of the datasets employed for IMG methodologies. Regrettably, public datasets for volume meshes are currently non-existent. Thus, our focus here is limited to commonly utilized datasets for triangle and quadrilateral meshes.

Triangle mesh. These commonly employed triangle mesh datasets include Princeton ModelNet¹, ShapeNet [150], TOSCA [151], COSEG [152], surface reconstruction benchmark of Berger [153] and Williams [67], Thingi10K [154], D-FAUST [155], Famous [80], and the CAD dataset ABC [156]; a dataset for single image 3D shape modeling: Pix3d [157]; facial expression datasets: COMA [59] and MeIn3D [158]; human body shapes datasets: MGN [11] and MultiHuman [159]; clothed body meshes with real texture: RenderPeople [160], Axyz [161], and Digit Wardrobe [11]; and indoor scenes datasets: ScanNet [162], Scenenet [163], Matterport3d [164],

1. <http://modelnet.cs.princeton.edu/>

TABLE 2: Challenges, advantages and limitations of all selected IMG methods. AMC denotes average monthly citations.

Article	Challenges	Advantages	Limitations	AMC
Self-organizing [42]	Nonuniform density mesh generation	Generates meshes with a specified density distribution	Only simple cases with given boundary are handled	0.12
Lowther <i>et al.</i> [43]	Nonuniform density mesh generation	Totally automatic, density adaptive mesh	Results extremely irregular, boundary needed	0.08
Alfonzetti <i>et al.</i> [44]	Nonuniform density mesh generation	Density adaptive mesh; boundary preserving	Require an initial rough mesh	0.12
Alfonzetti <i>et al.</i> [45]	Nonuniform density mesh generation	The vertex density function can be adaptively calculated	Require an initial rough mesh	0.07
Alfonzetti <i>et al.</i> [46]	Nonuniform tetrahedral mesh generation	Can generate a preset number of tetrahedral meshes	A rough initial mesh and density function are needed	0.04
Peng <i>et al.</i> [47]	Handle and represent the complex 3D object	Combine multilayer perception with 3D Object Reconstruction	The MLP only needs to recover the Z coordinate	0.14
Yao <i>et al.</i> [48]	Formulate mesh elements extraction rules for IMG	Reduces the number of singularities to a certain extent	Not end-to-end; may fall into a local minimum	0.16
Alfonzetti <i>et al.</i> [49]	Generate high quality meshes for 3D cases.	Good quality of the output meshes; simplicity of the algorithm	Only simple cases with given boundary are handled	0.04
Li <i>et al.</i> [50]	Robust to incomplete input data	Incomplete points cloud data surface reconstruction	Neural networks are only used to complement point clouds	0.02
Agostinho <i>et al.</i> [51]	Detailed multiresolution mesh generation	In a multiresolution fashion; only a simple initial mesh are needed	Only applicable to genus 0 object mesh generation	0.25
Wen <i>et al.</i> [52]	Ill-conditioned matrix and overfitting problem in RBF method	Overcome the numerical ill-conditioning of coefficient matrix and overfitting problem	The selection of the radial basis function limits the solution space of the implicit function	0.05
Xiong <i>et al.</i> [53]	Overcome the limitations of multistage processing	Jointly optimize geometry and connectivity; robust to outliers and noise; preserve sharp features	Convergence is not guaranteed; local minimum cannot be avoided; may fail in the hole regions	0.69
DeepGarment [8]	Intelligent, efficient and practical IMG	Efficient capture of 3D garment shapes from a single image only.	Only simple T-shirts and dresses can be handled.	0.89
Deepsketch2face [9]	Convert sketches to meshes efficiently	infer meshes from 2D sketches; effectively help users create face meshes.	Only for caricature model; Cannot create details such as wrinkles;	1.74
ShapeMVD [54]	Efficiently convert sketches to meshes	Output meshes preserve topology and shape structure	Output meshes lack details; not end-to-end differentiable	2.47
Surfnet [55]	Directly generates meshes of rigid and nonrigid shapes	Solves the problem of area distortion to a certain extent.	Only genus 0 surfaces can be generated	2.60
Pixel2mesh [10]	Output meshes lose surface details	Certainly maintains surface details	Only object similar to the initial mesh topology can be reconstructed	16.82
AtlasNet [56]	Generate high-resolution, 3D shapes	Generate shapes of arbitrary resolution; broad application	Exist distortion or overlap	16.29
Li <i>et al.</i> [57]	Modeling generic freeform 3D surfaces from sparse, expressive 2D sketches.	A freeform surface modeling; just needs succinct line annotations; solves the flow field by CNN	Cannot create geometric details; Model one patch each time; not end-to-end differentiable	1.22
DMC [58]	End-to-end marching cubes algorithm	End-to-end watertight meshes generation	High memory requirements and restricted to 32^3 voxel resolution	3.04
CoMA [59]	Generate 3D meshes of human faces	Generate diverse realistic 3D faces	Require a reference template from the same object class	6.60
3D-CFCN [60]	Incomplete and noise input with occlusions	To a certain extent, robust to noise	Poor mesh generation for thin slices	0.50
MGN [11]	Predict object geometry	Mesh reconstruction with texture captured from images	Cannot handle pose dependent deformations; rely on segmentation	5.03
3DN [35]	Simple and efficient IMG algorithm	An end-to-end optimizable model	Require a reference template with the same topology	2.23
TMN [12]	High quality mesh generation	A wide variety of topological meshes can be reconstructed	Unable to directly generate open surfaces	3.12
ONet [61]	Simple and efficient mesh representation	Can extract 3D meshes at any resolution; can handle various inputs	High memory requirements for 3D voxel grid	28.79
N3DMM [62]	Efficient feature representation for mesh	Create anisotropic mesh; lightweight and easy-to-optimize model	Require a reference template from the same object class	2.48
PGAN [63]	Explore geometric images for mesh generation	Generate diverse, realistic, 3D faces	Complex pre/post-process; mesh distortion by parameterization	0.03
HumanMeshNet [13]	Real-time reconstructions	Ensuring smooth surface reconstruction	Training data and test data must be the same basic model	0.35
DISN [64]	High-quality and detailed mesh generation	Capture fine-grained details and generate high-quality 3D models	Only able to handle objects with a clear background	8.45
IM-Net [65]	Generate high visual quality mesh	Efficient representation with coordinate information	Many nonsurface points yield many invalid calculations	17.10
Scan2Mesh [66]	Adaptation to incomplete scans	Cleaner and more CAD-like meshes from noisy and partial range scans	High memory requirements; does not enforce mesh regularity or continuity	1.67
DGP [67]	Dense reconstruction of the input point cloud	Robust to noise; capture sharp features; arbitrary resolution meshes.	High computational complexity; no adaptive patch selection	2.59
Mesh R-CNN [14]	High-quality and detailed mesh generation	Applicable to unconstrained real-world images with many objects	Nonmanifold, low reconstruction accuracy	8.03
DeepSDF [68]	Trade-offs across fidelity, efficiency and compression capabilities	Enables high-quality shape representation, interpolation and completion from partial and noisy 3D input data	Simple fully connected network architecture without local information or translation equivariance	32.23
Pixel2mesh++ [15]	Fusion of information across views	Exploit cross-view information; Strong generalization	Only genus 0 surface can be generated	3.74
PQ-Net [69]	Encoding of local structure and local geometry	Learn a 3D shape representation in the form of sequential part assembly	Cannot learn symmetry relations or produce topology-altering meshed	2.15
BCNet [16]	Accuracy of geometric representation	Different topology garments can be generated	Some details cannot be generated	2.48
PolyGen [70]	Generate a wide variety of realistic geometries	Is capable of generating coherent and diverse mesh samples	It is suitable for the generation of artificial object surfaces but not for the generation of complex surfaces.	3.68
DGTS [71]	Learn the geometric information of the grid	Transferring geometric information between two surfaces with different topologies; no parameterization needed	Inability to effectively learn the overall semantic information; high requirements for input data	1.30
Neural Subdivision [72]	Generate smooth and feature-preserving subdivision results.	Complex nonlinear subdivision schemes can be learned	Global semantic information cannot be guaranteed	1.83
Mobile3drecon [73]	Real-time mesh generation	Real-time dense mesh reconstruction	Poorly maintained sharp features	3.33
Sal [74]	Applicable to unsigned geometric data	Not needed for ground truth normal data or inside/outside labeling	Bad generation for thin parts	7.30
Pixel2mesh2 [17]	Preserve surface details	Geometric regularization constrains	Only for genus 0 surface	0.76
Voxel2mesh [40]	No postprocessing needed	Can generate 3D meshes without any postprocessing	Just suitable for genus 0 surface generation	1.39
X-ray2shape [18]	Applicable to low-contrast image data	Effective for low-contrast image as input	Too many singularities; cannot handle complex topology	0.30
Li <i>et al.</i> [19]	Background reconstruction	The final mesh less affected by the environment	Preprocessing determines mesh quality; thin part is poorly generated	0.02
Meshlet [75]	Robust to sparse and noisy, 3D points	Valid for sparse and noisy data; regular and globally consistent meshes	Failure on thin structures; the resolution is fixed	1.08
LIG [76]	IMG for scale and complexity indoor scenes	Can generate meshes for arbitrary scale scenes with details	Assume that different categories share similar part geometries	7.85
CONet [77]	Scalable and complex scenes mesh generation	Integrate local and global information; obtain translation equivariance	Not applicable to sparse point cloud and not robust to noise	13.59
ILSM [78]	Realistic liquid splashes mesh generation	Generate realistic liquid splash meshes from simple user sketch input	Unable to generate detailed splatters as real-world splatters; weak generalization	0.22
IER [79]	Generate fine-grained details; good generalizability	Preserve details; handle ambiguous structures; good generalizability	Nonmanifold mesh with holes; need posttreatment for holes	0.83
Yang <i>et al.</i> [20]	Generate object details	End-to-end architecture	Only applicable to genus 0 shape	0
MesherNet [21]	Automatic unstructured mesh generation.	A nonuniform mesh generation method; Combine PDEs with IMG	Only planar mesh generation was tested	0.81
Point2Surf [80]	Handle partial scan and noise.	Applicable to nonuniform noisy input, objects with varying topological	Not an end-to-end differentiable process	2.73
DEFTET [81]	Single image to tetrahedral meshes generation	Applicable to any topology; Efficient algorithm; no postprocessing	With flipped tetrahedrons; feature not guaranteed	1.38
BTM [36]	Generalizability of single mesh prototype	More generalizable than comparable methods	High computational complexity, multiple mesh prototypes needed	0.10
PointTriNet [82]	Efficient, scalable and generalizable algorithms	Trained in an unsupervised manner, robustness to outliers	Mesh with many holes, nonmanifold	1.04
Surface Hof [22]	High-resolution surface reconstruction	Generate detailed mesh at arbitrary resolution for various topologies	Only for images with a clear background; not end-to-end differentiable	0.10

TABLE 2: Challenges, advantages and limitations of all selected IMG methods. AMC denotes average monthly citations.

Article	Challenges	Advantages	Limitations	AMC
REIN [83]	Mesh generation with sparse point cloud	Mesh generation for sparse point clouds,	Nonmanifold, no permutation invariance	0.04
Henderson <i>et al.</i> [23]	Produce textured meshes	No self-intersections; does not need ground-truth segmentation masks	Require a reference template with the same topology	1.83
SSRNet [84]	Large-scale point clouds mesh reconstruction	Strong scalability; good at reconstructing geometry details	Not an end-to-end differentiable process; the partition method is crucial	1.26
MeshVAE [85]	An effective pooling operation of mesh	Pooling operation based on mesh simplification; can generate details	Only for uniform mesh; fail with nonwatertight or irregular mesh	0.65
Point2Mesh [37]	Learnable priori for mesh generation	Robust to unoriented normals and noise; produce watertight meshes; unsupervised learning	Initial mesh needed; high computational complexity and low reconstruction accuracy	3.68
NMF [38]	Manifold mesh generation	Generate two-manifold meshes	Only for genus 0 shapes	1.26
NDF [86]	High-resolution outputs of arbitrary shape	Can generate open surface; can represent inner structures	Need postprocessing for meshing	3.67
Meshsdf [87]	A differentiable way to produce surface mesh	End-to-end differentiable; watertight mesh for arbitrary topology	Insufficient generalization performance, only generate simple meshes	2.19
Smirnov <i>et al.</i> [26]	Efficient IMG.	Final mesh with fewer singularities	Postprocessing needed	0.87
Lasr [24]	Nonrigid structures mesh generation	Does not need category-specific mesh template; good generalization	Fails at heavy occlusions; efficiency needs improvement	1.47
Transformerfusion [88]	Efficient coding and accurate reconstruction	Online reconstruction approach running at interactive frame-rates	Cannot construct details for some scenes	2.36
CSPNet [89]	Efficient and low memory representation for complex surfaces	Can represent complex surfaces of any topology; efficient computation of local geometric properties	Not an end-to-end differentiable process	0.82
DASM [25]	Utilization of local regularization	A plug-and-play smoothing module can generate smoother meshes.	Hand-designed metric needed for mesh smooth	0.47
Neuralrecon [90]	Accuracy, Consistency, and Real-time reconstruction	Can generator accurate and coherent reconstruction in real-time	Not an end-to-end differentiable process	2.93
Sa-convoonet [91]	Scalability to large-scale scenes; Generalizability	Valid for large scenes	Slow inference speed	1.18
DMTet [92]	High-quality and detailed mesh from IMG	Can generate meshes with arbitrary topology, finer geometric details, fewer artifacts	Global uniform resolution; suffers from bad local minimum; tends to produce double/broken surfaces	0.81
Vis2mesh [93]	Large-scale scenes mesh generation	Good generalization; robust to noise; detail reconstruction	The system is complex	0.09
IMLSNet [94]	Transform the discrete point sets into smooth mesh	Define implicit functions on point sets	Not an end-to-end differentiable process	1.60
Retrievalfuse [95]	Mesh reconstruction of large scenes	Accurate scene reconstructions; maintain local detail	The results of the retrieval may be suboptimal	1.27
Deepdt [96]	Inside/outside classification cannot obtain clean meshes	Without ground truth labels of tetrahedrons or visibility information	Unable to handle excessively large input	0.47
Iso-points [97]	IMG for noisy and incomplete input	Faster convergence and accurate recovery of details and topology	Not explicitly model the appearance	0.87
DST [98]	The differentiable structure of triangular mesh	Control both the vertex positions and the topology; linear complexity to the number of vertices	Surface partition needed; visible artifacts exist across boundaries; cannot handle numerous points or patches	0.38
SAP [99]	Efficient and differentiable point to mesh layer	Output watertight manifold meshes; interpretable, lightweight and short inference time; robustness to outliers	Limited to small scenes; cubic memory requirements	1.60
Bertiche <i>et al.</i> [27]	Generalizability to various geometry and topology	Allows continuous predictions with differential geometry	Missing garments details in the final mesh	0
NeeDrop [100]	Mesh generation with extremely sparse point cloud	A self-supervised method; input data can be sparse point clouds	Requires postprocessing to generate meshes	0.33
AnalyticMesh [101]	Accurate mesh reconstruction	Does not lose the details of the implicit field compared to the Marching Cubes	postprocessing, such as smoothing or holes filling are needed	0
LMR [28]	Mesh inference with in-the-wild videos	Local-dynamical-modeling approach to video mesh recovery	Generated mesh cannot match the human body	0
NRSfM [29]	Nonrigid surface reconstruction	A simple and efficient model for nonrigid deformed mesh generation	The generated mesh may overlap at patch boundaries	0.11
DHSP [102]	Mesh generation with sparse colored point cloud; multipriiori integration	Generate mesh with high-resolution texture; robust to sparse and noisy point clouds	Complex structure and operation; only single object models with simple topology can be processed	0.27
Neural-Pull [39]	Learning high-quality SDFs to generate mesh	Learn SDF without the signed distance value; robust to noise	Missing sharp features in the reconstruction mesh	1.30
DI-Fusion [103]	The tradeoff between memory and mesh quality	Simultaneously encode geometric and uncertain information	Unable to preserve spatial continuity sharp features	1.47
DSE [104]	Generate manifold meshes	Produces near-manifold triangulations, robustness to outliers	Not E2E differentiable; require alignment for patches	0.67
LDFQ [105]	Most of the IMG methods of quadrilateral rely on a dense user-provided direction field	A robust data-driven approach to the computation of direction fields	The way the ground truth of the frame field is constructed determines the quality of the mesh	0.09
DGNN [106]	IMG with Large-scale scenes and incomplete point cloud input	Mesh reconstruction for large-scale, defect-laden point clouds; take into account visibility information	Low generalization, limit precision when the acquisition is noisy, missing reconstruction details	0.17
Hu <i>et al.</i> [30]	Generation of complex geometric features	Achieve a trade-off between mesh density and feature representation	The final mesh has more overlap	0.07
NMC [107]	IMG method that persevere sharp features	Can maintain sharp geometric features and learning local topology	Sensitive to rotation; possible self-intersection in the generated mesh	0.89
Skeletonnet [32]	Lack of constraints on the topology.	Topology-preserved; high-quality skeletal volume	Cannot deal with natural images in the wild; high algorithm complexity	0.80
Nvdifrec [108]	Joint optimization of topology, material and lighting	An appearance-aware and end-to-end mesh generator with materials.	High calculation and memory consumption; simplified shading model	3.33
Selfrecon [33]	Combine the advantages of implicit and explicit representations	Produce high-fidelity surfaces for arbitrarily clothed humans with self-supervised optimization	Long optimization time; mainly suitable for self-rotating motions	2.00
Autosdf [109]	Powerful priors for mesh generation	Useful for multimodal generation for a diverse set of tasks	Sensitive to alignment; CAD models only; not end-to-end differentiable	3.67
NKF [110]	Generate large scenes mesh with sparse point cloud; generalizability	Can reconstruct shape categories outside the training set; can reconstruct large scenes mesh	Kernel implementation requires a dense linear solve; requirement of oriented points	2
Sketch2PQ [111]	Real-time IMG	Real-time mesh generation with a dense direction field	Not E2E differentiable; only for disk topology surfaces	0
SRMAE [112]	Good generalizability	Can handle meshes in different sizes; suitable for unbounded mesh	Cumbersome preprocessing; underutilization of global information	0.25
OnSurfacePrior [113]	IMG with sparse point cloud and valid prior	SDF and normals are not needed; applicable to sparse point clouds	Not an end-to-end differentiable process	1.33
Lu <i>et al.</i> [114]	Intelligent advancing-front method;	An automation advancing front based IMG method	Valid only for 2D mesh; no control over singularities	0
MGNNet [115]	Differential structured mesh generation	Unsupervised structured quad mesh generation; requires no prior knowledge or measured dataset	Boundary accuracy determines the quality of the mesh; only for planar mesh	0.33
RLQMG [116]	Efficient and intelligent quad-mesh generation.	Automatic quad-mesh generation without extra clean-up operations	Lack of control over singularities; only for planar mesh generation	0.20
TopoNet [34]	Topological generalizability	Not limited by the topology of the template; better geometry capture	The reconstructed meshes with small holes	0
SAniHead [41]	Mesh generation of animal heads	Can generate meshes with geometric details	Hard to create shapes with thin structures and poor generalization ability	1.33
NDC [117]	Difficult to obtain surface gradients for Dual Contouring method	Good reconstruction quality; applicable to various inputs	Produce nonmanifold mesh, not completely invariant to orientation	0
PCGAN [118]	Process topology similarity among meshes by CNN	Preserve topology information and the spatial structure	Need preprocess and postprocess for mesh generation	0
Nice-slam [119]	Large scenes oversmooth mesh generation	Can fill small holes and extrapolate scene geometry; real-time system	Coarse representation only	4.33
NRGBD [120]	Scenes high-quality mesh reconstruction of room-scale scenes	Effective combination of depth observations and neural radiance field for room-scale scene reconstruction	Need postprocessing; lack of details; slow convergence	8.67
POCO [121]	Scalability of implicit reconstruction	Suitable for both single-object reconstruction and whole-scene reconstruction	Cannot complete shapes when missing large parts; without normals, yield orientation failures	2.00

and Synthetic Rooms [77]. The two most frequently used datasets are ShapeNet and ModelNet. ShapeNet contains 3D models from a multitude of semantic categories and organizes them under the Word-Net taxonomy. ShapeNet has indexed more than 3,000,000 3D models, 220,000 of which are classified into 3,135 categories. ModelNet has 662 object categories and 127915 CAD models. It contains three subsets: Modelnet10 with 10 categories; Modelnet40 with 40 categories; and Aligned40 with 40 classes aligned 3D model. **Quadrilateral mesh.** The commonly used quadrilateral mesh datasets include QuadWild [165]; Clothed body meshes with real texture: RenderPeople [160], Axyz [161]. Notably, the quadrilateral mesh datasets are not sufficient, which also hinders the development of this field to a large extent.

7.2 Commonly Used Prior

In IMG, commonly employed underlying priors include *smoothness*, *primitives*, *distribution*, *user-driven*, *orientation*, *visibility* and *regularity* priors. In the absence of supervision information, these priors provide effective guidance for model optimization. Below, we introduce them in detail.

The surface *smoothness* prior constrains the reconstructed surface to satisfy a certain level of smoothness. The most general form is local smoothness, e.g., the Laplace smooth constraint, which often serves as a regular term of the objective function [10], [14]–[17], [20], [25], [40], [81]. Geometric regularization [111], [113] attempts to pull points to their neighbor. With smooth prior constraints, the model often filters out noise but also loses some high-frequency details.

The geometric *primitive* prior assumes that the scene geometry may be explained by a compact set of simple geometric shapes, i.e., planes, boxes, spheres, and cylinders.

The *orientation* prior is an important prior in traditional mesh generation and is applicable to deep learning-based mesh generation tasks. The normal consistency prior guarantees the consistency of orientation when generating object surfaces based on distance fields [74].

The *visibility* prior makes assumptions about the exterior space of the reconstructed scene and how this can provide cues for combating noise, nonuniform sampling, and missing data. Specifically, Deng *et al.* used *visibility* in Sketch2PQ [111] by a depth-normal compatibility term and depth sample term. Sulzer *et al.* [106] skillfully utilized visibility information derived from camera positions and produced watertight meshes. Song *et al.* [93] explicitly employs depth completion for the visibility prediction of 3D points.

The global *regularity* prior takes advantage of the notion that many shapes possess a certain level of regularity in their higher-level composition. Regularity refers to the relationship between the whole and the part [69], [76]. The same kind of object is composed of similar basic subparts in a similar arrangement. For example, cars are composed of wheels, chassis, and cover from bottom to top. In addition, symmetry is also a common regularity.

7.3 Classical mesh-based learning architectures

With the success of deep learning methods in computer vision, many neural network models have been introduced to conduct 3D shape representation using volumetric grids [68], [166]–[169] and point clouds [170]–[176]. However, due

to the complexity and irregularity of mesh data, there are only a few neural networks that simultaneously use the geometric and topological information as inputs [177]–[180]. MeshCNN [177], a convolutional neural network designed for triangular meshes, defined edge-centric convolution and pooling operations. Convolutions are applied on edges and the four edges of their incident triangles, and pooling is applied via an edge collapse operation that retains the surface topology. These operations facilitate a direct analysis of mesh in its native form. MeshNet [178] defined face-centric convolution and pooling operations for triangular meshes. To synthesize geometric information and topological information, the network provides an effective data pre-processing and network framework. PD-MeshNet [179] utilized features for both edges and faces of a 3D mesh as input and dynamically aggregated them using an attention mechanism. Singh [180] designed surface correlation blocks that capture local features at various scales.

Although only a few of the network models mentioned above can simultaneously process geometry and topology, many outstanding works have emerged in the research of mesh analysis and feature representation. Convolutions that exploit mesh connectivity structures were introduced by Masci *et al.* [181]. Since then, several approaches have addressed the irregularity of the mesh structure and sampling rate, proposing ideas such as uniformly sampling the neighborhood of each vertex or achieving regularity by spectral decomposition [182]–[186]. Recently, combined with a learned heat diffusion operation, DiffusionNet [187] offered a unified perspective across representations of surface geometry. Another line of work seeks to understand the structure and connectivity of the mesh [38], [85], [185]. The authors implemented convolutional layers over point features related to vertices, for example, in EdgeConv [174]. By combining neighborhood or edge information, these methods can achieve a better comprehensive representation of the geometric and topological information of the mesh surface.

8 DISCUSSION AND CONCLUSION

IMG significantly enhances the generalizability, robustness, and practical utility of traditional mesh generation techniques. However, it comes with its own set of challenges: dealing with non-ideal data, algorithm practicality, and complex object generation. The non-ideal data challenge entails working with 2D images lacking 3D information, sparse and non-uniform point clouds, occluded or incomplete point clouds, noisy data, and large-scale input data. Algorithm practicality presents hurdles such as ensuring broad generalization, achieving high computational efficiency, maintaining low memory requirements, exhibiting robustness against non-ideal inputs, and guaranteeing end-to-end differentiability. The challenge of generating meshes for complex objects pertains to intricate objects or scenes. This includes the creation of meshes for large and dynamic scenes, high density detail, texture based meshes, water splash meshes, non-uniform density meshes, and structured meshes.

8.1 Summary of open challenges

In this review, we have highlighted the limitations found across the selected articles. These limitations not only represent the challenges within the IMG field, but also help

identify potential future research directions. The primary challenges we've recognized are as follows:

(1) **Generalization:** Generalization means the ability for an algorithm to handle objects not observed during training - is a crucial aspect demonstrating the practical utility of the algorithm. The lack of generalization is a prevalent issue with current IMG methods, particularly those using deformation-based mesh generation techniques. These methods often produce meshes specific to a particular object type or topology, limiting their wider applicability.

(2) **Interpretability:** The interpretability of IMG is critically important, especially within the aerospace industry. It is our suggestion that melding IMG with fundamental mathematical and physical models, as evidenced by studies such as Zheng et al. [188] and Lei et al. [189], can not only effectively govern model optimization but also impart enhanced interpretability to the model. However, thus far, no research has directly addressed this particular issue.

(3) **Lack of Benchmark:** As discussed in Section 6.2, performing a comprehensive comparative analysis of all these IMG methods poses significant challenges. The lack of standardized testing procedures, coupled with variations in test data and metrics, has significantly impeded IMG's advancement. Therefore, the establishment of benchmarking standards would substantially expedite the progress of IMG.

(4) **Watertight and manifold:** Watertight and manifold meshes are frequently required in many computer graphics applications. However, current element classification based IMG methods cannot generate watertight or manifold meshes. In attempting to generate a mesh from scratch, these methods overlook certain local constraints. Furthermore, due to the discrete nature of watertight and manifold meshes, they fail to provide gradients essential for model optimization.

(5) **Structured mesh:** Structured meshes offer numerous advantages that unstructured meshes cannot match, including highly efficient storage and access. Unfortunately, due to their inherent complexity, MGNet [115] is currently the only IMG method that considers structured mesh generation. Thus, it is critical to explore more practical structured or semi-structured IMG algorithms capable of handling real-world objects. Additionally, there is a noticeable scarcity of quadrilateral and volume mesh datasets, which hinders the advancement of IMG in areas of structured quadrilateral mesh and volume mesh generation. Creating open-source datasets in these fields could significantly expedite the development of IMG in the context of structured meshes.

(6) **Textured mesh:** Textured mesh is a highly sought-after representation for 3D objects and finds extensive application in numerous fields, including industrial design and digital entertainment. Existing IMG research has primarily concentrated on the challenges of reconstructing geometry and topology. In contrast, only a handful of studies have focused on the generation of textured meshes [23], [108], [190]. Furthermore, the generation of textured meshes undeniably enhances the interpretability of the model, making this an important area for future exploration in the IMG field.

(7) **Large or dynamic scene mesh generation:** Scene mesh generation holds immense potential in diverse fields, such as autonomous driving, indoor robotics, and mixed reality. For IMG methods to truly prove useful in real-world applications, several critical properties must be addressed. Firstly,

real-time functionality of the algorithm is desired. Secondly, the algorithm should have the capacity to make plausible predictions for regions devoid of observations. Furthermore, the system should demonstrate scalability to accommodate large scenes. Lastly, the robustness against noisy or incomplete observations is essential. Existing methodologies [13], [28], [73], [90], [91], [93], [95], [110], [119], however, fall short of satisfying these criteria simultaneously, resulting in scene meshes that tend to be overly smooth and lacking in detail.

8.2 Conclusion

In this article, we have provided a systematic and comprehensive review of Intelligent Mesh Generation (IMG) methodologies, highlighting core techniques, application scopes, learning goals, data types, targeted challenges, strengths, and limitations. We scrutinized 113 research papers for a detailed analysis and data extraction. Our primary points of extraction included 1) the addressed challenges, 2) the fundamental concepts, application range, advantages, and disadvantages, 3) the type of input data, the nature of output mesh, and mesh quality, and 4) potential directions for future research. Articles were classified based on their techniques, unit elements, and applicable data types. While there have been substantial advancements in IMG in recent years, we also identified numerous issues and challenges that pave the way for future exploration and investigation in this field.

To the best of our knowledge, this is the most recent and comprehensive survey of existing IMG methods. This survey provides a holistic perspective and extensive research resources for scholars in the field of IMG. However, it is worth noting that our review has certain limitations. Specifically, our focus was solely on IMG; non-machine-learning mesh generation methods were beyond our purview. As a result, numerous valuable papers on traditional mesh generation were not included. Furthermore, despite the application of systematic literature review methodology and a manual search to ensure comprehensive inclusion, there may have been relevant papers that were overlooked in the initial selection due to the sheer volume of existing literature.

ACKNOWLEDGMENT

This research was supported by the National Key R&D Program of China (2021YFA1003003) and the National Natural Science Foundation of China under Grants 61936002 and T2225012.

REFERENCES

- [1] J. P. Slotnick, A. Khodadoust, J. Alonso, D. Darmofal, W. Gropp, E. Lurie, and D. J. Mavriplis, "Cfd vision 2030 study: a path to revolutionary computational aerosciences," Tech. Rep., 2014.
- [2] S. Keele, "Guidelines for performing systematic literature reviews in software engineering," 2007.
- [3] M. Berger, A. Tagliasacchi, L. M. Seversky, P. Alliez, G. Guennebaud, J. A. Levine, A. Sharf, and C. T. Silva, "A survey of surface reconstruction from point clouds," in *Comput Graph Forum*, vol. 36, no. 1, 2017, pp. 301–329.
- [4] X.-F. Han, H. Laga, and M. Bennamoun, "Image-based 3d object reconstruction: State-of-the-art and trends in the deep learning era," *IEEE TPAMI*, vol. 43, no. 5, pp. 1578–1604, 2019.
- [5] D. Khan, A. Plopski, Y. Fujimoto, M. Kanbara, G. Jabeen, Y. J. Zhang, X. Zhang, and H. Kato, "Surface remeshing: A systematic literature review of methods and research directions," *IEEE TVCG*, vol. 28, no. 3, pp. 1680–1713, 2020.

- [6] Y.-P. Xiao, Y.-K. Lai, F.-L. Zhang, C. Li, and L. Gao, "A survey on deep geometry learning: From a representation perspective," *Comput Vis Media*, vol. 6, no. 2, p. 8, 2020.
- [7] W. Nianhua, L. Peng, C. Xinghua, and Z. Laiping, "Preliminary investigation on unstructured mesh generation technique based on advancing front method and machine learning methods," *J. Theor. Appl. Mech.*, vol. 53, no. 3, pp. 740–751, 2021.
- [8] R. Daněrek, E. Dibra, C. Öztireli, R. Ziegler, and M. Gross, "DeepGarment : 3D Garment Shape Estimation from a Single Image," *Comput Graph Forum*, vol. 36, pp. 269–280, May 2017.
- [9] X. Han, C. Gao, and Y. Yu, "DeepSketch2face: a deep learning based sketching system for 3d face and caricature modeling," *ACM TOG*, vol. 36, no. 4, pp. 1–12, 2017.
- [10] N. Wang, Y. Zhang, Z. Li, Y. Fu, W. Liu, and Y.-G. Jiang, "Pixel2mesh: Generating 3d mesh models from single rgb images," in *ECCV*, 2018, pp. 52–67.
- [11] B. L. Bhatnagar, G. Tiwari, C. Theobalt, and G. Pons-Moll, "Multi-garment net: Learning to dress 3d people from images," in *ICCV*, 2019, pp. 5420–5430.
- [12] J. Pan, X. Han, W. Chen, J. Tang, and K. Jia, "Deep mesh reconstruction from single rgb images via topology modification networks," in *ICCV*, 2019, pp. 9964–9973.
- [13] A. Venkat, C. Patel, Y. Agrawal, and A. Sharma, "HumanMesh-Net: Polygonal Mesh Recovery of Humans," in *ICCVW*, 2019, pp. 2178–2187.
- [14] G. Gkioxari, J. Johnson, and J. Malik, "Mesh R-CNN," in *ICCV*, 2019, pp. 9784–9794.
- [15] C. Wen, Y. Zhang, Z. Li, and Y. Fu, "Pixel2mesh++: Multi-view 3d mesh generation via deformation," in *ICCV*, 2019, pp. 1042–1051.
- [16] B. Jiang, J. Zhang, Y. Hong, J. Luo, L. Liu, and H. Bao, "Bcnet: Learning body and cloth shape from a single image," in *ECCV*, 2020, pp. 18–35.
- [17] N. Wang, Y. Zhang, Z. Li, Y. Fu, H. Yu, W. Liu, X. Xue, and Y.-G. Jiang, "Pixel2mesh: 3d mesh model generation via image guided deformation," *IEEE TPAMI*, vol. 43, no. 10, pp. 3600–3613, 2020.
- [18] F. Tong, M. Nakao, S. Wu, M. Nakamura, and T. Matsuda, "X-ray2shape: reconstruction of 3d liver shape from a single 2d projection image," in *2020 42nd EMBC*, 2020, pp. 1608–1611.
- [19] X. Li, K. Ping, X. Gu, and M. He, "3d shape reconstruction of furniture object from a single real indoor image," in *2020 17th ICCWAMTIP*, 2020, pp. 101–104.
- [20] Y. Dongsheng, K. Ping, and X. Gu, "3D Reconstruction based on GAT from a Single Image," in *ICCWAMTIP*, 2020, pp. 122–125.
- [21] Z. Zhang, Y. Wang, P. K. Jimack, and H. Wang, "MeshingNet: A New Mesh Generation Method Based on Deep Learning," in *ICCS*, 2020, pp. 186–198.
- [22] Z. Wang, V. Isler, and D. D. Lee, "Surface Hof: Surface Reconstruction From A Single Image Using Higher Order Function Networks," in *ICIP*, 2020, pp. 2666–2670.
- [23] P. Henderson, V. Tsiminaki, and C. H. Lampert, "Leveraging 2D Data to Learn Textured 3D Mesh Generation," in *CVPR*, Jun. 2020, pp. 7495–7504.
- [24] G. Yang, D. Sun, V. Jampani, D. Vlasic, F. Cole, H. Chang, D. Ramanan, W. T. Freeman, and C. Liu, "Lasr: Learning articulated shape reconstruction from a monocular video," in *CVPR*, 2021, pp. 15980–15989.
- [25] U. Wickramasinghe, P. Fua, and G. Knott, "Deep active surface models," in *CVPR*, 2021, pp. 11 652–11 661.
- [26] D. Smirnov, M. Bessmeltsev, and J. Solomon, "Learning manifold patch-based representations of man-made shapes," in *ICLR*, 2020.
- [27] H. Bertiche, M. Madadi, and S. Escalera, "Deep Parametric Surfaces for 3D Outfit Reconstruction from Single View Image," in *FG 2021*, 2021, pp. 1–8.
- [28] R. Li, S. Karanam, R. Li, T. Chen, B. Bhanu, and Z. Wu, "Learning Local Recurrent Models for Human Mesh Recovery," in *3DV*, 2021, pp. 555–564.
- [29] M. Pedone, A. Mostafa, and J. Heikkilä, "Learning non-rigid surface reconstruction from spatio-temporal image patches," in *ICPR*, 2021, pp. 10 134–10 140.
- [30] M. Hu, J. Rutqvist, and C. I. Steefel, "Mesh generation and optimization from digital rock fractures based on neural style transfer," *J. Rock Mech. Geotech. Eng.*, vol. 13, Aug. 2021.
- [31] N. Wang, Y. Zhang, Z. Li, Y. Fu, H. Yu, W. Liu, X. Xue, and Y.-G. Jiang, "Pixel2Mesh: 3D Mesh Model Generation via Image Guided Deformation," *IEEE TPAMI*, vol. 43, pp. 3600–3613, 2021.
- [32] J. Tang, X. Han, M. Tan, X. Tong, and K. Jia, "Skeletonnet: A topology-preserving solution for learning mesh reconstruction of object surfaces from rgb images," *IEEE TPAMI*, 2021.
- [33] B. Jiang, Y. Hong, H. Bao, and J. Zhang, "Selfrecon: Self reconstruction your digital avatar from monocular video," in *CVPR*, 2022, pp. 5605–5615.
- [34] T. Ben Charrada, H. Tabia, A. Chetouani, and H. Laga, "TopoNet: Topology Learning for 3D Reconstruction of Objects of Arbitrary Genus," *Comput Graph Forum*, no. n/a, Mar. 2022.
- [35] W. Wang, D. Ceylan, R. Mech, and U. Neumann, "3dn: 3d deformation network," in *CVPR*, 2019, pp. 1038–1046.
- [36] T. Rios, J. Kong, B. van Stein, T. Bäck, P. Wollstadt, B. Sendhoff, and S. Menzel, "Back To Meshes: Optimal Simulation-ready Mesh Prototypes For Autoencoder-based 3D Car Point Clouds," in *SSCI*, 2020, pp. 942–949.
- [37] R. Hanocka, G. Metzger, R. Giryes, and D. Cohen-Or, "Point2Mesh: A Self-Prior for Deformable Meshes," *ACM TOG*, vol. 39, no. 4, Aug. 2020.
- [38] K. Gupta and M. Chandraker, "Neural mesh flow: 3D manifold mesh generation via diffeomorphic flows," in *NeurIPS*, Dec. 2020, pp. 1747–1758.
- [39] B. Ma, Z. Han, Y.-S. Liu, and M. Zwicker, "Neural-Pull: Learning Signed Distance Functions from Point Clouds by Learning to Pull Space onto Surfaces," *ICML*, vol. 139, Jan. 2021.
- [40] U. Wickramasinghe, E. Remelli, G. Knott, and P. Fua, "Voxel2mesh: 3d mesh model generation from volumetric data," in *MICCAI*, 2020, pp. 299–308.
- [41] D. Du, X. Han, H. Fu, F. Wu, Y. Yu, S. Cui, and L. Liu, "SAni-Head: Sketching Animal-Like 3D Character Heads Using a View-Surface Collaborative Mesh Generative Network," *IEEE TVCG*, vol. 28, no. 6, pp. 2415–2429, Jun. 2022.
- [42] A. Chang-Hoi, L. Sang-Soo, L. Hyuek-Jae, and L. Soo-Young, "A self-organizing neural network approach for automatic mesh generation," *IEEE Trans. Magn.*, vol. 27, pp. 4201–4204, Sep. 1991.
- [43] D. Lowther and D. Dyck, "A density driven mesh generator guided by a neural network," *IEEE Trans. Magn.*, vol. 29, no. 2, pp. 1927–1930, Mar. 1993.
- [44] S. Alfonzetti, S. Coco, S. Cavalieri, and M. Malgeri, "Automatic mesh generation by the let-it-grow neural network," *IEEE Trans. Magn.*, vol. 32, no. 3, pp. 1349–1352, 1996.
- [45] S. Alfonzetti, "A finite element mesh generator based on an adaptive neural network," *IEEE Trans. Magn.*, vol. 34, no. 5, pp. 3363–3366, Sep. 1998.
- [46] ——, "A neural network generator for tetrahedral meshes," *IEEE Trans. Magn.*, vol. 39, no. 3, pp. 1650–1653, 2003.
- [47] L. W. Peng and S. M. Shamsuddin, "3D object reconstruction and representation using neural networks," in *GRAPHITE*, 2004, pp. 139–147.
- [48] S. Yao, B. Yan, B. Chen, and Y. Zeng, "An ann-based element extraction method for automatic mesh generation," *Expert Syst. Appl.*, vol. 29, no. 1, pp. 193–206, 2005.
- [49] S. Alfonzetti, E. Dilettoso, and N. Salerno, "An Optimized Generator of Finite Element Meshes Based on a Neural Network," *IEEE Trans. Magn.*, vol. 44, pp. 1278–1281, Jun. 2008.
- [50] G.-x. Li, X.-m. Wu, and W.-m. Zhao, "Incomplete Points Cloud Data Surface Reconstruction Based on Neural Network," in *IIH-MSP*, 2008, pp. 913–916.
- [51] A. de Medeiros Brito Junior, A. D. DÓria Neto, J. Dantas de Melo, and L. M. Garcia Goncalves, "An Adaptive Learning Approach for 3-D Surface Reconstruction From Point Clouds," *IEEE Trans. Neural Netw.*, vol. 19, pp. 1130–1140, 2008.
- [52] P. Wen, X. Wu, Y. Zhu, and X. Peng, "LS-RBF network based 3D surface reconstruction method," in *CCDC*, Jun. 2009, pp. 5785–5789, iSSN: 1948-9447.
- [53] S. Xiong, J. Zhang, J. Zheng, J. Cai, and L. Liu, "Robust surface reconstruction via dictionary learning," *ACM TOG*, vol. 33, no. 6, pp. 201:1–201:12, Nov. 2014.
- [54] Z. Lun, M. Gadelha, E. Kalogerakis, S. Maji, and R. Wang, "3d shape reconstruction from sketches via multi-view convolutional networks," in *3DV*, 2017, pp. 67–77.
- [55] A. Sinha, A. Unmesh, Q. Huang, and K. Ramani, "Surfnets: Generating 3d shape surfaces using deep residual networks," in *CVPR*, 2017, pp. 6040–6049.
- [56] T. Groueix, M. Fisher, V. G. Kim, B. C. Russell, and M. Aubry, "A papier-mâché approach to learning 3d surface generation," in *CVPR*, 2018, pp. 216–224.

- [57] C. Li, H. Pan, Y. Liu, X. Tong, A. Sheffer, and W. Wang, "Robust flow-guided neural prediction for sketch-based freeform surface modeling," *ACM TOG*, vol. 37, no. 6, pp. 1–12, 2018.
- [58] Y. Liao, S. Donné, and A. Geiger, "Deep Marching Cubes: Learning Explicit Surface Representations," in *CVPR*, Jun. 2018, pp. 2916–2925.
- [59] A. Ranjan, T. Bolkart, S. Sanyal, and M. J. Black, "Generating 3d faces using convolutional mesh autoencoders," in *ECCV*, 2018, pp. 704–720.
- [60] Y.-P. Cao, Z.-N. Liu, Z.-F. Kuang, L. Kobbelt, and S.-M. Hu, "Learning to reconstruct high-quality 3d shapes with cascaded fully convolutional networks," in *ECCV*, 2018, pp. 616–633.
- [61] L. Mescheder, M. Oechsle, M. Niemeyer, S. Nowozin, and A. Geiger, "Occupancy networks: Learning 3d reconstruction in function space," in *CVPR*, 2019, pp. 4460–4470.
- [62] G. Bouritsas, S. Bokhnyak, S. Ploumpis, M. Bronstein, and S. Zafeiriou, "Neural 3d morphable models: Spiral convolutional networks for 3d shape representation learning and generation," in *ICCV*, 2019, pp. 7213–7222.
- [63] T. Li, Y. Shi, X. Sun, J. Wang, and B. Yin, "PGAN: Prediction Generative Adversarial Nets for Meshes," in *VCIP*, 2019, pp. 1–4.
- [64] Q. Xu, W. Wang, D. Ceylan, R. Mech, and U. Neumann, "DISN: Deep Implicit Surface Network for High-quality Single-view 3D Reconstruction," in *NeurIPS*, vol. 32, 2019.
- [65] Z. Chen and H. Zhang, "Learning Implicit Fields for Generative Shape Modeling," in *CVPR*, Jun. 2019, pp. 5932–5941.
- [66] A. Dai and M. Nießner, "Scan2Mesh: From Unstructured Range Scans to 3D Meshes," in *CVPR*, Jun. 2019, pp. 5569–5578.
- [67] F. Williams, T. Schneider, C. Silva, D. Zorin, J. Bruna, and D. Panozzo, "Deep Geometric Prior for Surface Reconstruction," in *CVPR*, Jun. 2019, pp. 10122–10131.
- [68] J. J. Park, P. Florence, J. Straub, R. Newcombe, and S. Lovegrove, "DeepSDF: Learning continuous signed distance functions for shape representation," in *CVPR*, 2019, pp. 165–174.
- [69] R. Wu, Y. Zhuang, K. Xu, H. Zhang, and B. Chen, "Pq-net: A generative part seq2seq network for 3d shapes," in *CVPR*, 2020, pp. 829–838.
- [70] C. Nash, Y. Ganin, S. M. A. Eslami, and P. Battaglia, "PolyGen: An autoregressive generative model of 3D meshes," in *ICML*, vol. 119. PMLR, 2020, pp. 7220–7229.
- [71] A. Hertz, R. Hanocka, R. Giryes, and D. Cohen-Or, "Deep geometric texture synthesis," *ACM TOG*, vol. 39, pp. 108–1, 2020.
- [72] H.-T. D. Liu, V. G. Kim, S. Chaudhuri, N. Aigerman, and A. Jacobson, "Neural subdivision," *ACM TOG*, vol. 39, pp. 124–1, 2020.
- [73] X. Yang, L. Zhou, H. Jiang, Z. Tang, Y. Wang, H. Bao, and G. Zhang, "Mobile3drecon: real-time monocular 3d reconstruction on a mobile phone," *IEEE TVCG*, vol. 26, no. 12, pp. 3446–3456, 2020.
- [74] M. Atzmon and Y. Lipman, "Sal: Sign agnostic learning of shapes from raw data," in *CVPR*, 2020, pp. 2565–2574.
- [75] A. Badki, O. Gallo, J. Kautz, and P. Sen, "Meshlet priors for 3d mesh reconstruction," in *CVPR*, 2020, pp. 2849–2858.
- [76] C. Jiang, A. Sud, A. Makadia, J. Huang, M. Nießner, T. Funkhouser *et al.*, "Local implicit grid representations for 3d scenes," in *CVPR*, 2020, pp. 6001–6010.
- [77] S. Peng, M. Niemeyer, L. Mescheder, M. Pollefeys, and A. Geiger, "Convolutional occupancy networks," in *ECCV*, 2020, pp. 523–540.
- [78] G. Yan, Z. Chen, J. Yang, and H. Wang, "Interactive liquid splash modeling by user sketches," *TOG*, vol. 39, pp. 1–13, 2020.
- [79] M. Liu, X. Zhang, and H. Su, "Meshing Point Clouds with Predicted Intrinsic-Extrinsic Ratio Guidance," in *ECCV*. Springer International Publishing, 2020, pp. 68–84.
- [80] P. Erler, P. Guerrero, S. Ohrhallinger, N. J. Mitra, and M. Wimmer, "Points2Surf Learning Implicit Surfaces from Point Clouds," in *ECCV*, 2020, pp. 108–124.
- [81] J. Gao, W. Chen, T. Xiang, A. Jacobson, M. McGuire, and S. Fidler, "Learning Deformable Tetrahedral Meshes for 3D Reconstruction," in *NeurIPS*, vol. 33, 2020, pp. 9936–9947.
- [82] N. Sharp and M. Ovsjanikov, "PointTriNet: Learned Triangulation of 3D Point Sets," in *ECCV*, 2020, pp. 762–778.
- [83] R. Daroya, R. Atienza, and R. Cajote, "REIN: Flexible Mesh Generation from Point Clouds," in *CVPRW*, Jun. 2020, pp. 1444–1453.
- [84] Z. Mi, Y. Luo, and W. Tao, "SSRNet: Scalable 3D Surface Reconstruction Network," in *CVPR*, Jun. 2020, pp. 967–976.
- [85] Y.-J. Yuan, Y.-K. Lai, J. Yang, Q. Duan, H. Fu, and L. Gao, "Mesh Variational Autoencoders with Edge Contraction Pooling," in *CVPRW*, Jun. 2020, pp. 1105–1112.
- [86] J. Chibane, G. Pons-Moll *et al.*, "Neural unsigned distance fields for implicit function learning," *NeurIPS*, vol. 33, pp. 21638–21652, 2020.
- [87] E. Remelli, A. Lukoianov, S. Richter, B. Guillard, T. Bagautdinov, P. Baque, and P. Fua, "Meshsdf: Differentiable iso-surface extraction," *NeurIPS*, vol. 33, pp. 22468–22478, 2020.
- [88] A. Bozic, P. Palafox, J. Thies, A. Dai, and M. Nießner, "Transformerfusion: Monocular rgbd scene reconstruction using transformers," *NeurIPS*, vol. 34, pp. 1403–1414, 2021.
- [89] R. Venkatesh, T. Karmali, S. Sharma, A. Ghosh, R. V. Babu, L. A. Jeni, and M. Singh, "Deep implicit surface point prediction networks," in *ICCV*, 2021, pp. 12653–12662.
- [90] J. Sun, Y. Xie, L. Chen, X. Zhou, and H. Bao, "Neuralrecon: Real-time coherent 3d reconstruction from monocular video," in *CVPR*, 2021, pp. 15598–15607.
- [91] J. Tang, J. Lei, D. Xu, F. Ma, K. Jia, and L. Zhang, "Sa-cononet: Sign-agnostic optimization of convolutional occupancy networks," in *ICCV*, 2021, pp. 6504–6513.
- [92] T. Shen, J. Gao, K. Yin, M.-Y. Liu, and S. Fidler, "Deep marching tetrahedra: a hybrid representation for high-resolution 3d shape synthesis," in *NeurIPS*, vol. 34, 2021, pp. 6087–6101.
- [93] S. Song, Z. Cui, and R. Qin, "Vis2mesh: Efficient mesh reconstruction from unstructured point clouds of large scenes with learned virtual view visibility," in *ICCV*, 2021, pp. 6514–6524.
- [94] S.-L. Liu, H.-X. Guo, H. Pan, P.-S. Wang, X. Tong, and Y. Liu, "Deep implicit moving least-squares functions for 3d reconstruction," in *CVPR*, 2021, pp. 1788–1797.
- [95] Y. Siddiqui, J. Thies, F. Ma, Q. Shan, M. Nießner, and A. Dai, "Retrievalfuse: Neural 3d scene reconstruction with a database," in *ICCV*, 2021, pp. 12568–12577.
- [96] Y. Luo, Z. Mi, and W. Tao, "Deepdpt: Learning geometry from delaunay triangulation for surface reconstruction," in *AAAI*, vol. 35, no. 3, 2021, pp. 2277–2285.
- [97] W. Yifan, S. Wu, C. Oztureli, and O. Sorkine-Hornung, "Iso-points: Optimizing neural implicit surfaces with hybrid representations," in *CVPR*, 2021, pp. 374–383.
- [98] M.-J. Rakotosaona, N. Aigerman, N. J. Mitra, M. Ovsjanikov, and P. Guerrero, "Differentiable surface triangulation," *ACM TOG*, vol. 40, no. 6, pp. 1–13, 2021.
- [99] S. Peng, C. Jiang, Y. Liao, M. Niemeyer, M. Pollefeys, and A. Geiger, "Shape as points: A differentiable poisson solver," *NeurIPS*, vol. 34, pp. 13032–13044, 2021.
- [100] A. Boulch, P.-A. Langlois, G. Puy, and R. Marlet, "NeeDrop: Self-supervised Shape Representation from Sparse Point Clouds using Needle Dropping," in *3DV*, 2021, pp. 940–950.
- [101] J. Lei, K. Jia, and Y. Ma, "Learning and Meshing from Deep Implicit Surface Networks Using an Efficient Implementation of Analytic Marching," *IEEE TPAMI*, pp. 1–1, 2021.
- [102] X. Wei, Z. Chen, Y. Fu, Z. Cui, and Y. Zhang, "Deep Hybrid Self-Prior for Full 3D Mesh Generation," *ICCV*, 2021.
- [103] J. Huang, S.-S. Huang, H. Song, and S.-M. Hu, "DI-Fusion: Online Implicit 3D Reconstruction with Deep Priors," in *CVPR*, Jun. 2021, pp. 8928–8937.
- [104] M.-J. Rakotosaona, P. Guerrero, N. Aigerman, N. Mitra, and M. Ovsjanikov, "Learning Delaunay Surface Elements for Mesh Reconstruction," in *CVPR*, Jun. 2021, pp. 22–31.
- [105] A. Dielen, I. Lim, M. Lyon, and L. Kobbelt, "Learning Direction Fields for Quad Mesh Generation," *Comput Graph Forum*, vol. 40, no. 5, pp. 181–191, Aug. 2021.
- [106] R. Sulzer, L. Landrieu, R. Marlet, and B. Vallet, "Scalable Surface Reconstruction with Delaunay-Graph Neural Networks," *Comput Graph Forum*, vol. 40, no. 5, pp. 157–167, Aug. 2021.
- [107] Z. Chen and H. Zhang, "Neural Marching Cubes," *ACM TOG*, vol. 40, no. 6, pp. 1–15, Dec. 2021.
- [108] J. Munkberg, J. Hasselgren, T. Shen, J. Gao, W. Chen, A. Evans, T. Müller, and S. Fidler, "Extracting triangular 3d models, materials, and lighting from images," in *CVPR*, 2022, pp. 8280–8290.
- [109] P. Mittal, Y.-C. Cheng, M. Singh, and S. Tulsiani, "Autosdf: Shape priors for 3d completion, reconstruction and generation," in *CVPR*, 2022, pp. 306–315.
- [110] F. Williams, Z. Gojcic, S. Khamis, D. Zorin, J. Bruna, S. Fidler, and O. Litany, "Neural fields as learnable kernels for 3d reconstruction," in *CVPR*, 2022, pp. 18500–18510.

- [111] Z. Deng, Y. Liu, H. Pan, W. Jabi, J. Zhang, and B. Deng, "Sketch2PQ: Freeform Planar Quadrilateral Mesh Design via a Single Sketch," *IEEE TVCG*, pp. 1–1, 2022.
- [112] S. Hahner and J. Garcke, "Mesh Convolutional Autoencoder for Semi-Regular Meshes of Different Sizes," in *WACV*, 2022, pp. 2344–2353.
- [113] B. Ma, Y.-S. Liu, and Z. Han, "Reconstructing surfaces for sparse point clouds with on-surface priors," in *CVPR*, 2022, pp. 6315–6325.
- [114] P. Lu, N. Wang, Y. Lin, X. Zhang, Y. Wu, and H. Zhang, "A new unstructured hybrid mesh generation method based on BP-ANN," *J. Phys. Conf. Ser.*, 2022.
- [115] X. Chen, T. Li, Q. Wan, X. He, C. Gong, Y. Pang, and J. Liu, "MGNet: a novel differential mesh generation method based on unsupervised neural networks," *Eng. Comput.*, Mar. 2022.
- [116] J. Pan, J. Huang, G. Cheng, and Y. Zeng, "Reinforcement learning for automatic quadrilateral mesh generation: a soft actor-critic approach," Mar. 2022.
- [117] Z. Chen, A. Tagliasacchi, T. Funkhouser, and H. Zhang, "Neural Dual Contouring," *ACM TOG*, vol. 41, pp. 1–13, Jul. 2022.
- [118] T. Li, Y. Shi, X. Sun, J. Wang, and B. Yin, "PCGAN: Prediction-Compensation Generative Adversarial Network for Meshes," *IEEE T-CSVT*, vol. 32, no. 7, pp. 4667–4679, Jul. 2022.
- [119] Z. Zhu, S. Peng, V. Larsson, W. Xu, H. Bao, Z. Cui, M. R. Oswald, and M. Pollefeys, "Nice-slam: Neural implicit scalable encoding for slam," in *CVPR*, 2022, pp. 12786–12796.
- [120] D. Azinović, R. Martin-Brualla, D. B. Goldman, M. Nießner, and J. Thies, "Neural rgb-d surface reconstruction," in *CVPR*, 2022, pp. 6290–6301.
- [121] A. Boulch and R. Marlet, "Poco: Point convolution for surface reconstruction," in *CVPR*, 2022, pp. 6302–6314.
- [122] W. E. Lorensen and H. E. Cline, "Marching cubes: A high resolution 3d surface construction algorithm," *ACM siggraph computer graphics*, vol. 21, no. 4, pp. 163–169, 1987.
- [123] A. Doi and A. Koide, "An efficient method of triangulating equi-valued surfaces by using tetrahedral cells," *IEICE Trans. Inf. Syst.*, vol. 74, pp. 214–224, 1991.
- [124] M. Kazhdan and H. Hoppe, "Screened poisson surface reconstruction," *ACM TOG*, vol. 32, no. 3, pp. 1–13, 2013.
- [125] P. Wang, L. Liu, Y. Liu, C. Theobalt, T. Komura, and W. Wang, "Neus: Learning neural implicit surfaces by volume rendering for multi-view reconstruction," in *NeurIPS*.
- [126] J. Wang, T. Bleja, and L. Agapito, "Go-surf: Neural feature grid optimization for fast, high-fidelity rgb-d surface reconstruction," in *3DV*, 2022, pp. 433–442.
- [127] S. J. Ruuth and B. Merriman, "A simple embedding method for solving partial differential equations on surfaces," *J Comput Phys*, vol. 227, no. 3, pp. 1943–1961, 2008.
- [128] C. B. Macdonald and S. J. Ruuth, "Level set equations on surfaces via the closest point method," *J. Sci. Comput.*, vol. 35, pp. 219–240, 2008.
- [129] ——, "The implicit closest point method for the numerical solution of partial differential equations on surfaces," *SIAM Journal on Scientific Computing*, vol. 31, no. 6, pp. 4330–4350, 2010.
- [130] T. Märtz and C. B. Macdonald, "Calculus on surfaces with general closest point functions," *SINUM*, vol. 50, pp. 3303–3328, 2012.
- [131] L. Tian, C. B. Macdonald, and S. J. Ruuth, "Segmentation on surfaces with the closest point method," in *ICIP*, 2009, pp. 3009–3012.
- [132] C. B. Macdonald, J. Brandman, and S. J. Ruuth, "Solving eigenvalue problems on curved surfaces using the closest point method," *J Comput Phys*, vol. 230, no. 22, pp. 7944–7956, 2011.
- [133] D.-T. Lee and B. J. Schachter, "Two algorithms for constructing a delaunay triangulation," *In. j. comput. inf. sci.*, vol. 9, no. 3, pp. 219–242, 1980.
- [134] N. Ray, W. C. Li, B. Lévy, A. Sheffer, and P. Alliez, "Periodic global parameterization," *ACM TOG*, vol. 25, no. 4, pp. 1460–1485, 2006.
- [135] X. Gu, S. J. Gortler, and H. Hoppe, "Geometry images," in *Proceedings of the 29th annual conference on Computer graphics and interactive techniques*, 2002, pp. 355–361.
- [136] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, pp. 139–144, 2020.
- [137] R. Löhner and P. Parikh, "Generation of three-dimensional unstructured grids by the advancing-front method," *Int. J. Numer. Methods Fluids*, vol. 8, no. 10, pp. 1135–1149, 1988.
- [138] Y. Guo, X. Huang, Z. Ma, Y. Hai, R. Zhao, and K. Sun, "An improved advancing-front-delaunay method for triangular mesh generation," in *CGI*, 2021, pp. 477–487.
- [139] T. Ju, F. Losasso, S. Schaefer, and J. Warren, "Dual contouring of hermite data," in *SIGGRAPH*, 2002, pp. 339–346.
- [140] D. P. Kingma and M. Welling, "Auto-encoding variational bayes," *arXiv preprint arXiv:1312.6114*, 2013.
- [141] H. Fan, H. Su, and L. J. Guibas, "A point set generation network for 3d object reconstruction from a single image," in *CVPR*, 2017, pp. 605–613.
- [142] O. Sorkine, D. Cohen-Or, and S. Toledo, "High-pass quantization for mesh encoding," in *SGP*, vol. 42, 2003, p. 3.
- [143] I. Abouelaziz, M. El Hassouni, and H. Cherifi, "A curvature based method for blind mesh visual quality assessment using a general regression neural network," in *SITIS*. IEEE, 2016, pp. 793–797.
- [144] I. Abouelaziz, M. E. Hassouni, and H. Cherifi, "A convolutional neural network framework for blind mesh visual quality assessment," in *ICIP*, Sep. 2017, pp. 755–759.
- [145] I. Abouelaziz, A. Chetouani, M. El Hassouni, and H. Cherifi, "Reduced Reference Mesh Visual Quality Assessment Based on Convolutional Neural Network," in *SITIS*, 2018, pp. 617–620.
- [146] I. Abouelaziz, A. Chetouani, M. E. Hassouni, L. J. Latecki, and H. Cherifi, "Convolutional Neural Networks for Blind Mesh Visual Quality Assessment Using 3D Visual Saliency," in *ICIP*, 2018.
- [147] I. Abouelaziz, A. Chetouani, M. El Hassouni, H. Cherifi, and L. J. Latecki, "Learning graph convolutional network for blind mesh visual quality assessment," *IEEE Access*, vol. 9, pp. 108 200–108 211, 2021.
- [148] X. Chen, J. Liu, Y. Pang, J. Chen, L. Chi, and C. Gong, "Developing a new mesh quality evaluation method based on convolutional neural network," *Eng. Appl. Comput. Fluid Mech*, vol. 14, no. 1, pp. 391–400, 2020.
- [149] X. Chen, J. Liu, C. Gong, S. Li, Y. Pang, and B. Chen, "Mve-net: An automatic 3-d structured mesh validity evaluation framework using deep neural networks," *Comput. Aided Des.*, vol. 141, p. 103104, 2021.
- [150] A. X. Chang, T. Funkhouser, L. Guibas, P. Hanrahan, Q. Huang, Z. Li, S. Savarese, M. Savva, S. Song, H. Su *et al.*, "Shapenet: An information-rich 3d model repository," *arXiv preprint arXiv:1512.00675*, 2015.
- [151] A. M. Bronstein, M. M. Bronstein, and R. Kimmel, "Efficient computation of isometry-invariant distances between surfaces," *SIAM J. Sci. Comput.*, vol. 28, pp. 1812–1836, 2006.
- [152] Y. Wang, S. Asafi, O. Van Kaick, H. Zhang, D. Cohen-Or, and B. Chen, "Active co-analysis of a set of shapes," *ACM TOG*, vol. 31, no. 6, pp. 1–10, 2012.
- [153] M. Berger, J. A. Levine, L. G. Nonato, G. Taubin, and C. T. Silva, "A benchmark for surface reconstruction," *ACM TOG*, vol. 32, no. 2, pp. 1–17, 2013.
- [154] Q. Zhou and A. Jacobson, "Thingi10k: A dataset of 10,000 3d-printing models," *arXiv preprint arXiv:1605.04797*, 2016.
- [155] F. Bogo, J. Romero, G. Pons-Moll, and M. J. Black, "Dynamic faust: Registering human bodies in motion," in *CVPR*, 2017, pp. 6233–6242.
- [156] S. Koch, A. Matveev, Z. Jiang, F. Williams, A. Artemov, E. Burnaev, M. Alexa, D. Zorin, and D. Panozzo, "Abc: A big cad model dataset for geometric deep learning," in *CVPR*, 2019, pp. 9601–9611.
- [157] X. Sun, J. Wu, X. Zhang, Z. Zhang, C. Zhang, T. Xue, J. B. Tenenbaum, and W. T. Freeman, "Pix3d: Dataset and methods for single-image 3d shape modeling," in *CVPR*, 2018, pp. 2974–2983.
- [158] J. Booth, A. Roussos, S. Zafeiriou, A. Ponniah, and D. Dunaway, "A 3d morphable model learnt from 10,000 faces," in *CVPR*, 2016, pp. 5543–5552.
- [159] Y. Zheng, R. Shao, Y. Zhang, T. Yu, Z. Zheng, Q. Dai, and Y. Liu, "Deepmulticap: Performance capture of multiple characters using sparse multiview cameras," in *ICCV*, 2021, pp. 6239–6249.
- [160] "Renderpeople," <https://renderpeople.com/free-3d-people>.
- [161] "axyz," <https://secure.axyz-design.com/>, 2019.
- [162] A. Dai, A. X. Chang, M. Savva, M. Halber, T. Funkhouser, and M. Nießner, "Scannet: Richly-annotated 3d reconstructions of indoor scenes," in *CVPR*, 2017, pp. 5828–5839.
- [163] A. Handa, V. Pătrăucean, S. Stent, and R. Cipolla, "Scenenet: An annotated model generator for indoor scene understanding," in *ICRA*, 2016, pp. 5737–5743.

- [164] A. Chang, A. Dai, T. Funkhouser, M. Halber, M. Niessner, M. Savva, S. Song, A. Zeng, and Y. Zhang, "Matterport3d: Learning from rgb-d data in indoor environments," *arXiv preprint*, 2017.
- [165] N. Pietroni, S. Nuvoli, T. Alderighi, P. Cignoni, and M. Tarini, "Reliable feature-line driven quad-remeshing," in *ACM SIGGRAPH*, vol. 40, no. 4, 2021, pp. 1–17.
- [166] D. Maturana and S. Scherer, "Voxnet: A 3d convolutional neural network for real-time object recognition," in *IROS*, 2015, pp. 922–928.
- [167] A. Brock, T. Lim, J. M. Ritchie, and N. Weston, "Generative and discriminative voxel modeling with convolutional neural networks," *arXiv preprint arXiv:1608.04236*, 2016.
- [168] P.-S. Wang, Y. Liu, Y.-X. Guo, C.-Y. Sun, and X. Tong, "Ocnn: Octree-based convolutional neural networks for 3d shape analysis," *ACM TOG*, vol. 36, pp. 1–11, 2017.
- [169] Z. Chen, V. G. Kim, M. Fisher, N. Aigerman, H. Zhang, and S. Chaudhuri, "Decor-gan: 3d shape detailization by conditional refinement," in *CVPR - ICS*, 2021, pp. 15735–15744.
- [170] C. R. Qi, H. Su, K. Mo, and L. J. Guibas, "Pointnet: Deep learning on point sets for 3d classification and segmentation," in *CVPR*, 2017, pp. 652–660.
- [171] C. R. Qi, L. Yi, H. Su, and L. J. Guibas, "Pointnet++: Deep hierarchical feature learning on point sets in a metric space," *NeurIPS*, vol. 30, 2017.
- [172] Y. Yang, C. Feng, Y. Shen, and D. Tian, "Foldingnet: Point cloud auto-encoder via deep grid deformation," in *CVPR*, 2018, pp. 206–215.
- [173] H. Thomas, C. R. Qi, J.-E. Deschaud, B. Marcotegui, F. Goulette, and L. J. Guibas, "Kpconv: Flexible and deformable convolution for point clouds," in *ICCV*, 2019, pp. 6411–6420.
- [174] Y. Wang, Y. Sun, Z. Liu, S. E. Sarma, M. M. Bronstein, and J. M. Solomon, "Dynamic graph cnn for learning on point clouds," *ACM Trans. Graph.*, vol. 38, no. 5, oct 2019.
- [175] Z. Li, W. Wang, N. Lei, and R. Wang, "Weakly supervised point cloud upsampling via optimal transport," in *ICASSP*, 2022, pp. 2564–2568.
- [176] R. Wiersma, A. Nasikun, E. Eisemann, and K. Hildebrandt, "Deltaconv: anisotropic operators for geometric deep learning on point clouds," *ACM TOG*, vol. 41, no. 4, pp. 1–10, 2022.
- [177] R. Hanocka, A. Hertz, N. Fish, R. Giryes, S. Fleishman, and D. Cohen-Or, "Meshcnn: a network with an edge," *ACM TOG*, vol. 38, no. 4, pp. 1–12, 2019.
- [178] Y. Feng, Y. Feng, H. You, X. Zhao, and Y. Gao, "Meshnet: Mesh neural network for 3d shape representation," in *AAAI*, vol. 33, no. 01, 2019, pp. 8279–8286.
- [179] F. Milano, A. Loquercio, A. Rosinol, D. Scaramuzza, and L. Carlone, "Primal-dual mesh convolutional neural networks," *NeurIPS*, vol. 33, pp. 952–963, 2020.
- [180] V. V. Singh, S. V. Sheshappanavar, and C. Kambhamettu, "Meshnet++: A network with a face," in *ACM MM*, 2021, pp. 4883–4891.
- [181] J. Masci, D. Boscaini, M. Bronstein, and P. Vandergheynst, "Geodesic convolutional neural networks on riemannian manifolds," in *ECCVW*, 2015, pp. 37–45.
- [182] D. Boscaini, J. Masci, E. Rodolà, and M. Bronstein, "Learning shape correspondence with anisotropic convolutional neural networks," *NeurIPS*, 2016.
- [183] I. Lim, A. Dielen, M. Campen, and L. Kobbelt, "A simple approach to intrinsic correspondence learning on unstructured 3d meshes," in *ECCVW*, 2018, pp. 0–0.
- [184] J. Schult, F. Engelmann, T. Kontogianni, and B. Leibe, "Dualconvmesh-net: Joint geodesic and euclidean convolutions on 3d meshes," in *CVPR*, 2020, pp. 8612–8622.
- [185] Y. Zhou, C. Wu, Z. Li, C. Cao, Y. Ye, J. Saragih, H. Li, and Y. Sheikh, "Fully convolutional mesh autoencoder using efficient spatially varying kernels," *NeurIPS*, vol. 33, pp. 9251–9262, 2020.
- [186] D. Smirnov and J. Solomon, "Hodgenet: Learning spectral geometry on triangle meshes," *ACM TOG*, vol. 40, pp. 1–11, 2021.
- [187] N. Sharp, S. Attaiki, K. Crane, and M. Ovsjanikov, "Diffusionnet: Discretization agnostic learning on surfaces," *ACM TOG*, vol. 41, no. 3, pp. 1–16, 2022.
- [188] X. Zheng, Y. Zhu, W. Chen, N. Lei, Z. Luo, and X. Gu, "Quadrilateral mesh generation III: Optimizing singularity configuration based on abel-jacobi theory," *Comput. Methods Appl. Mech. Eng.*, vol. 387, p. 114146, 2021.
- [189] N. Lei, K. Su, L. Cui, S.-T. Yau, and X. D. Gu, "A geometric view of optimal transportation and generative model," *Comput. Aided Geom. Des.*, vol. 68, pp. 1–21, 2019.
- [190] L. Gao, T. Wu, Y.-J. Yuan, M.-X. Lin, Y.-K. Lai, and H. Zhang, "Tm-net: Deep generative networks for textured meshes," *ACM TOG*, vol. 40, no. 6, pp. 1–15, 2021.



Na Lei received her B.S. degree in 1998 and a Ph.D. degree in 2002 from Jilin University. Currently, she is a professor at Dalian University of Technology. Her research interest is the application of modern differential geometry and algebraic geometry to solve problems in engineering and medical fields. She mainly focuses on computational conformal geometry, computer mathematics, and its applications in computer vision and geometric modeling.



Zezeng Li received a B.S. degree from Beijing University of Technology (BJUT) in 2015. He is currently pursuing a Ph.D. degree from Dalian University of Technology (DUT). His research interests include image processing, point cloud processing, and mesh generation.



Zebin Xu received a B.S. degree from Dalian University of Technology (DUT) in 2020. He is currently pursuing a Ph.D. degree from Dalian University of Technology (DUT). His research interests include image processing and mesh generation.



Ying Li received a B.S. degree in 1999 and a Ph.D. degree in 2004 from Jilin University. She completed her postdoctoral research at Tsinghua University in 2007. Currently, she is an associate professor at Jilin University. Her research interests are the application of machine learning models in computational biology and image processing.



Xianfeng Gu received a B.S. degree from Tsinghua University and a Ph.D. degree from Harvard University. He is now a tenured professor in the Department of Computer Science and Applied Mathematics at the State University of New York at Stony Brook. He has won several awards, such as the NSF CAREER Award of the USA, the Chinese Overseas Outstanding Youth Award,

the Chinese Fields Medal, and the Chenxing Golden Prize in Applied Mathematics. Professor Gu's team combines differential geometry, algebraic topology, Riemann surface theory, partial differential equations, and computer science to create a cross-disciplinary "computational conformal geometry", which is widely utilized in computer graphics, computer vision, 3D geometric modeling and visualization, wireless sensor networks, medical images, and other fields.