RRCNet: Rivet Region Classification Network for Rivet flush Measurement Based on 3D Point Cloud

Qian Xie, Dening Lu, Anyi Huang, Jianping Yang, Dawei Li, Yuan Zhang, Jun Wang

Abstract—In aircraft manufacturing industry, rivet inspection is a vital task for the aircraft structure stability and aerodynamic performance. In this paper, we propose a novel framework for fully automated rivet flush measurement, which is the key step in rivet inspection task. To efficiently perform rivet flush measurement, we first develop a mobile 3D scanning system to automatically capture the 3D point cloud of the aircraft skin surface. Subsequently, rivet regions are extracted through point cloud processing techniques. Instead of relying on handcrafted features, we propose a novel data-driven approach for rivet point extraction via a deep learning based technique. Our algorithm takes a scanned point cloud of the aircraft skin surface as input, and produces a dense point cloud label result for each point, distinguishing as rivet point or not. To achieve this, we propose a Rivet Region Classification Network (RRCNet) that can input 2D representations of a point and output a binary label indicating the point is rivet or non-rivet point. Moreover, we design a Field Attention Unit (FAU) to assign adaptive weights to different forms of 2D representations via the attention mechanism in convolutional neural networks. The extracted rivet regions can then be used to perform rivet flush measurement. The above components result in a fully automatic contactless measurement framework of aircraft skin rivet flush. Several experiments are performed to demonstrate the priority of the proposed RRCNet and the effectiveness of the presented rivet flush measurement framework.

Index Terms—3D deep learning, point cloud processing, rivet flush measurement, attention mechanism.

I. INTRODUCTION

Rivet has been widely used in aircraft manufacturing field, owning to its excellent characteristics as a permanent mechanical fastener. The riveting quality has a vital impact on the flight performance, especially for those high-speed aircrafts, as illustrated in Figure 1(d). Moreover, bad rivets may weaken the stealthy performance of stealth aircraft by increasing the Radar Cross Section (RCS). Thus, it is quite essential to inspect the quality of rivets on the aircraft skin surface.

Manuscript received August 14, 2020; revised September 11, 2020; accepted September 24, 2020. This work was supported in part by National Natural Science Foundation of China under Grant 61772267, the Fundamental Research Funds for the Central Universities under Grant NE2016004, and the NUAA Fundamental Research Funds under Grant NS2015053. (Corresponding author: Jun Wang.)

Qian Xie, Dening Lu, Jianping Yang, Dawei Li, Yuan Zhang and Jun Wang are with the College of Mechanical and Electrical Engineering, Nanjing University of Aeronautics and Astronautics, China (qianxie@nuaa.edu.cn; deninglu@gmail.com; jianpingyang90@163.com; dawei@nuaa.edu.cn; zhangyuan299@foxmail.com; wjun@nuaa.edu.cn).

Anyi Huang is with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, China (huanganyi@nuaa.edu.cn).

Rivet flush is the most important inspection index, which reflects the joint degree between the rivet and the aircraft skin surface. However, there exist two challenges to be solved in rivet flush measurement. First, there are too many rivets, over thousands, on one aircraft skin surface. It is a huge project to inspect all the rivets one by one, which, however, is essential since each rivet counts. Second, the quantitative measurement of rivet flush still remains an opening problem. In recent years, the 3D scanning technique is adopted to obtain the 3D shape information, e.g. point cloud, of the skin surface with rivets, as shown in Figure 1. Based on the captured 3D point cloud, rivet flush measurement can then be performed quantitatively [1]. Nevertheless, how to efficiently and accurately perform rivet flush measurement based on the captured 3D data is still a challenging problem.

To address the above two issues, we propose a fully automatic aircraft skin rivet inspection framework, which is composed of a data acquisition part and a data analysis part. To obtain the 3D point cloud on the surface with high efficiency, we develop a mobile scanning system with a 3D scanner, which can replace manual scanning procedure. The proposed scanning system is capable of capturing 3D points on the surface while moving following a pre-set path. After we obtain the scanned 3D point cloud, we need to extract the rivet region points before performing the following rivet flush measurement. Several methods have been proposed to perform rivet extraction, such as [1]. However, these method rely on merely one aspect of rivet structure, like density...
difference in [1]. Also, these previous methods are based on classical point cloud processing techniques which heavily rely on experimental parameters, thus lack of generalization. As far as we know, we are the first one to employ the deep learning technique in rivet structure detection from 3D point cloud.

In this paper, we present a deep learning based rivet extraction algorithm, which performs rivet point extraction by classifying each point as rivet or non-rivet point. Finding a good and suitable feature representation is quite important for the classification model [2]. Specifically, we first transfer the 3D point data into 2D representation via projecting 3D points into grid cells in 2D planes, resulting 2D maps encoding 3D geometric information. We then propose a Rivet Region Classification Network (RRCNet) to take the 2D maps as input and output the predictions in possibility of being rivet or not. In this way, the designed network can automatically learn the characteristics of rivet points with the given labelled data. Moreover, instead of using a single field (i.e., the height field) in most previous methods, we propose to use more kinds of fields to encode more geometric features and details in 2D maps. They are height, density, gaussian curvature and mean curvature fields. As far as we know, our method is the first one to extract rivets from point cloud using deep learning based technique.

In addition, we design a Field Attention Unit (FAU) in the proposed network to learn the most useful information in each field. We argue that some fields should contribute more to the final prediction. Nevertheless, the most important field map may differ from cases, making it difficult to select a fixed importance ranking for these fields for all cases. Hence, the proposed FAU leverages the attention mechanism to assign different weights to fields according to the encoded field information. The weighted information is then combined together to perform the final classification. In such way, our network can efficiently learn more useful information, which leads to improvements for the classification accuracy.

In rivet flush calculation stage, we need to first extract the rivet head points to measure the distance between the rivet head and the aircraft skin surface. However, the rivet contour points could be incorrectly segmented as rivet head points, which could bring incorrect flush values. To address this issue, we propose a hierarchical fitting method to extract the rivet head points. Rough rivet head regions are first extracted and then a further plane fitting step is performed to eliminate most outliers in the rivet contour.

Overall, the main contributions in this paper can be summarised as follows:

- We propose a fully automated rivet flush measurement framework which is capable of performing rivet flush measurement with high accuracy and efficiency.
- To automatically obtain the 3D shape information in the aircraft skin surface, we develop a flexible 3D auto-scanning system which can collect 3D point cloud by moving along a pre-set scan path without any user intervention.
- To extract rivet regions from the scanned 3D point cloud, we design a rivet region classification network (RRCNet), in which 2D maps of different fields are adaptively combined by a field attention unit (FAU).
- To get a more accurate rivet flush value, we introduce a hierarchical fitting strategy to exactly segment the rivet head points from the extracted rivet head region.

II. RELATED WORK

In this part, we divide our discussion on the related work into two parts, focusing on rivet flush measurement and 3D deep learning on point cloud.

A. Rivet Flush Measurement.

The quality of riveting is of great importance for the whole performance of the aircraft, making it quite necessary to perform rivet inspection in aviation industry [3], [4], [5]. Rivet inspection methods can be classified as two groups: 2D image-based and 3D point-based. 2D image-based methods take the image data of the aircraft skin as input. For instance, [6], [7], [8] proposed to perform rivet inspection and defect recognition based on the images captured by the magneto-optic imager (MOI). 2D images can also be applied to perform rivet detection, i.e., identification and location [9], [10]. Compared to 2D image based methods, 3D point data contains more geometric information, making it more suitable for geometry-related inspection tasks, such as rivet flush measurement. Paul et al. [11] proposed a perform rivet detection with the assistance of point cloud data captured with a RGB-D sensor, which makes their method be able to detect rivets in poorly illuminated environments. More recently, Xia et al. [12] used the fringe projection technique to obtain 3D point cloud data of aircraft surface. Based on the captured 3D data, they then performed rivet and seam structural defects detection. However, the rivet detection is still performed in 2D images in their method. Xie et al. [1] directly detected aircraft skin rivet on 3D point cloud via fitting rivet structure using multiple structures fitting technique. Thus, we employ 3D scanning technique to collect the shape information of the aircraft skin, and perform rivet inspection directly based on the captured 3D point cloud. However, instead of relying on handcrafted features as the above methods, we propose a data-driven approach to automatically learn the features of rivet pattern, inspired by the success of data-driven methods in various applications [13].

B. 3D deep learning on point cloud.

With the advances in 2D deep learning techniques, especially the convolutional neural networks (CNNs) [14], [15], [16], 3D deep learning based analysis on point cloud has become an important task in the field of computer vision and computer graphics [17], [18]. Current deep learning based methods on point cloud data can be grouped into 3 categories: Voxel-based methods [19], [20], [21], View-based methods and Point-based methods [22].

Voxel-based Methods. The biggest obstacle in 3D deep learning is the irregular characteristic of 3D point data. To address this issue, voxel-based methods transform the scattered
point cloud into regular structures, i.e., 3D volumetric grids, which can be processed conveniently by the 3D Convolution operation defined on regular 3D grids [23], [24]. Voxel-based methods have been successfully applied in various of applications, such as 3D object detection [25], object classification [26] and object segmentation [27], [28]. However, the performance of these methods heavily relies on the resolution of voxels, which makes them computationally demanding with large number of voxels. Another drawback is the inevitable geometric information loss during the quantization step of the voxelization procedure. Thus, these methods are unsuitable for tasks related to detailed feature perception.

**View-based Methods.** Considering the huge success of 2D CNNs on images, view-based methods [29], [30], [31] focus on projecting 3D point date onto 2D planes in multiple views, and then directly use 2D CNNs to conduct various analyses. View-based methods have achieved promising results at shape classification task. Nevertheless, these methods cannot be easily extended to tasks where local details are needed, such as point cloud consolidation [32].

**Point-based Methods.** Instead of transferring 3D points into other intermediate formats which could cause information loss, point-based approaches attempt to directly take the coordinates of 3D points as input. PointNet [33] and its extension PointNet++ [34] are the pioneering networks showing the capability of feature encoding using point-based methods. The basic idea behind the above two networks is to transfer each point as a vector by mapping 3D coordinates into higher dimensional space using multi-layer perceptrons (MLPs). Since the geometric characteristics of single points are determined by their neighboring points, Point-based approaches could distinguish more details than voxel-based approaches by relieving the quantization error problem caused by the voxelization operation. However, the adopted max-pooling operation for information aggregation in this kind of approaches makes them lack of the ability to distinguish the point-level details.

As can be seen, existing 3D deep learning methods on point cloud processing so far are focusing on coarse semantic understanding of object class label, such as object classification, object segmentation and object detection [35], [36]. Thus, it is hard to distinguish the tiny features and details in 3D shapes for the existing 3D deep learning techniques, which hinders their application in rivet point extraction.

We tackle this by combining the advantages of point-based and view-based approaches. Specifically, for each point, we get its neighborhood and encode the information of the neighborhood points by generating 2D projection maps. Moreover, we exploit the rivet structure in different fields. That is, we try to enlarge the difference between rivet and non-rivet regions by comparing them in several scalar value fields, such as height and curvature fields. We achieve this by first computing several scalar value fields on the scanned point cloud, and then encode these geometric information into 2D space by projecting 3D points in 2D planes. In such way, our method can efficiently learn the characteristics of rivet structure in multiple fields, and thus effectively classify points into rivet or non-rivet.

**III. OVERVIEW**

Figure 2 gives the overview of the proposed framework, which consists of two main phases, i.e., data acquisition and data processing. In the data acquisition phase, a mobile auto-scanning system is developed to capture the point cloud of the aircraft skin surface. In the data processing phase, a CNN-based rivet region extraction approach is presented to detect the rivet points in the scanned point cloud, followed by a rivet flush measurement method to perform the rivet quality assessment.

**Mobile scanning system.** We first develop a 3D scanning hardware system which contains a mobile platform, a robot and a 3D scanner. The mobile platform and robot can move flexibly following the scanning path order. Note that the 3D scanner can be switched according to different scanning tasks. 3D point cloud can be efficiently acquired with the 3D scanner moving over the aircraft skin surface.

**CNN-based rivet region extraction.** Taking as input the scanned point cloud, we extract the rivet region points via
a deep learning based classifier. Specifically, we propose a rivet region classification network (RRCNet) to classify points into rivet or non-rivet region. To improve the accuracy of the rivet extraction, we design a field attention unit (FAU) in the network to efficiently integrate different field information with the attention mechanism. In this way, we can obtain the group of individual rivet region point sets.

**Rivet flush measurement.** Each rivet region point set is finally fed into this stage to accomplish the rivet flush measurement. Rivet head points and non-rivet points are first divided by a RANSAC based plane fitting method. The fitted plane of the non-rivet points is then regarded as the reference plane to calculate the final flush values.

### IV. Mobile Scanning System

In this section, we focus on the design of the 3D scanning equipment to capture the point cloud of the aircraft skin surface. With the designed system, the operation of scanning can be fully automatic as the robot moves following a predetermined scanning path.

**3D Scanner.** With the development of 3D scanning technologies, various types of 3D scanners have been designed to collect 3D points on the object surface. The collected 3D points could be in huge number and of high precision, making it very suitable for quality inspection and shape measurement. After analyzing the requirement of the rivet flush measurement and the data quality of various types of 3D scanners, we choose Creaform MetraScan as our scanner. The precision of Creaform MetraScan can reach 0.025mm, and it can move around the object while scanning. A tracker is needed track the MetraScan to register scans into one unified coordinate system. Note that various 3D scanners such as Creaform Handyscan 700 or ATOS 3D scanner can also be easily integrated into the proposed system according to different measurement tasks. However, 3D scanners with low precision, such as Kinect or RealSense, is improper for our system, since they cannot get the detailed shape of riveting structures.

**Robot.** 3D scanner needs to move on the surface of objects to collect more data. Nevertheless, the manual hand movement is neither stable nor efficient. Hence, we adopt the collaborative robot to replace the function of the human arm. Here, we use the UR5 collaborative robot from Universal Robots. The weight and size of the UR5 robot is suitable for the auto-scanning system.

**Mobile platform.** In a fixed position, the space where the robot can reach is limited. When the object to be measured is too huge to scan in one position, we need to move the robot to another position to increase the reachable area. Thus, we place the robot in a mobile platform which can perform the movement automatically. The mobile platform also increase the flexibility of our auto-scanning system.

With the designed auto-scanning system, the 3D point cloud data collection can be totally automatic by following the path set in advance.

### V. CNN-based Rivet Region Extraction

After the point cloud with rivets are captured, we then extract rivet points by classifying every point into rivet or non-rivet categories. To this end, we first get the neighborhood point set and generate 2D maps for each point. These maps are regarded as the shape representation of point and then fed into RRCNet to perform classification.

#### A. Training data generation.

**2D maps generation.** Since the geometric information is encoded by the neighboring points, we use the neighborhood point set within a sphere centered at every point as the representation of the single point. The neighborhood point set here can be regarded as the local point patch (LPP). A LPP of the current point is first established \( r \)-radius searching in 3D space, we then use this patch to compute the rivet region probability. The \( r \)-radius of the neighborhood searching is set to be the 1.5x of the rivet radius. Inspired by [37], we transfer the 3D point data into 2D representation by projection operations. In this way, the 2D maps are capable of encoding the characteristics of rivet regions. Specifically, we first normalize every LPP by centering it and scaling it to unit size. Subsequently, a projecting plane \((p, v)\) is determined based on the vectors calculated by PCA analysis. Like [37], we divide the 2D planes into 56 \(\times\) 56 grids and then project all the \(r\)-radius neighboring points into these regular grids. The pixel values of these grids are determined by the scalar field values of the points.

Instead of merely generating one single 2D map, usually the height map, in most previous methods, we propose to use multiple fields to encode more geometric information. The four fields used in this paper are height, point density, Gaussian curvature and Mean curvature respectively.

- **Height map:** Height map is the most commonly used 2D representation for local point patch in previous methods. Given a point \( p_0 \), we first get its \( r\)-radius neighboring point set \( P = \{p_1, p_2, \ldots, p_n\} \). The vector \( v \) of the plane fitted by PCA analysis is the eigenvector corresponding to the smallest eigenvalue. The height \( h_i \) of the point \( p_i \in P \) to the plane can be then determined by:

  \[
  h_i = \frac{|v \cdot (p_i - p_0)|}{|v|}
  \]

  (1)

  Note that the original generated 2D height map contains several issues, such as grid containing no points. We thus
Field map generation

Field A

Field B

Field C

Field D

Feature extraction

CNN 1

CNN 1

CNN 1

CNN 1

Field pooling

FAU

FAU

FAU

FAU

Classification

CNN 2

CNN 2

CNN 2

CNN 2

Rivet region

Non-Rivet region

Fig. 4. Architecture of the proposed network, RRCNet, which consists of three main components: feature extraction part, field pooling part and classification part. RRCNet takes four field maps of a point as input and outputs the probability of being the rivet point.

The designed network can take more field maps as input to boost the classification performance further, as long as the added field maps can give more useful geometric information.

B. Network architecture.

Rivet Region Classification Network (RRCNet). The architecture of the proposed RRCNet is shown in Figure 4. It is comprised of three main components. The first component CNN1 is the backbone network for feature extraction, in which all the CNN blocks share the same weights. CNN1 consists of four convolutional layers, i.e., Conv1_1, Conv1_2, Conv2_1 and Conv2_2. Each two convolutional layers are followed by a max pooling layer to reduce the dimension of the input data. For each point in the given input (i.e., scanned 3D point cloud), we denote the set of projection 2D maps as

\[
\{f_1, f_2, f_3, f_4\}
\]

These 2D maps are then fed into CNN blocks to generate their corresponding feature maps

\[
F = \{f_1, f_2, f_3, f_4\}
\]

This is followed by the second component, a field pooling operation implemented by the proposed Field Attention Units (FAUs) with the shared parameters. The features are scaled by the weights generated by FAU, and then aggregated by the feature map concatenation operation. The resulting feature is finally processed by the third component, classification part, which contains CNN2 and a fully connected layer whose output represents the possibility of being rivet region or not. Details of the network architecture are given in Table I. Notably, batch normalization and a ReLU layer are followed with each convolution operation.

Field Attention Unit (FAU). In view-based learning methods, feature maps from different views are usually enhanced by assigning their corresponding relative importance before combined to perform the final classification. Inspired by this idea, we introduce a field attention unit (FAU) to predict the relative weight of each field based on the learned features, which is implemented by the attention mechanism [39], [40]. FAU intrinsically introduces dynamics conditioned on the field maps, which can be regarded as a self-attention function on

\[
D_p = \frac{N_p}{\frac{4}{3} \pi \cdot r_d^3}
\]

where \( \frac{4}{3} \pi \cdot r_d^3 \) is the 3D volume of the sphere with the given radius \( r_d \), and \( N_p \) is the number of points within the sphere. We set the radius to be 0.5x to the rivet head radius. Note that the neighborhood points searching of radius \( r_d \) here is different from the radius \( r \) for local point patch generation. \( r_d \) is used for local point density computation, and it is only \( \frac{r}{2} \) to \( r \).

Gaussian & Mean curvature map: Curvature is regarded as an important geometric property of surfaces in computer graphics. Although rivet head and non-rivet regions can be seen as planar, the curvature of rivet contour is different from the above two regions. Hence, curvature field is also included into our scalar field set. To encode as much geometric information as possible, we use two kinds of curvature, i.e., Gaussian curvature \( (K = k_1 \cdot k_2) \) and Mean curvature \( (H = 0.5 \cdot (k_1 + k_2)) \) [38], where \( k_1 \) and \( k_2 \) are the principal curvatures. Examples of the curvature fields can be seen in Figure 3.

Note that our input is not limited to these four field maps. The designed network can take more field maps as input to adopt the same post-processing schemes (e.g., Gaussian-weight interpolation) in [37] to deal with these problems. Please refer to [37] for more details. An example of the generated height maps can be found in Figure 3.

- Density map: Owning to the scanning mechanism and the geometric characteristic of the rivet structure, more points will be collected on the rivet contour region during scanning procedure. This results in the local point density difference between rivet and non-rivet regions, as shown in Figure 3. Therefore, the density field can be used as one of the maps to encode the geometric information of the rivet structure. For each point \( p \), its local point density \( D_p \) can be calculated as:

\[
D_p = \frac{N_p}{\frac{4}{3} \pi \cdot r_d^3}
\]
feature selection in the channel aspect. One can refer to [41] for more theoretical analysis on this kind of channel-wise attention mechanism.

Specifically, the architecture details of the proposed FAU are given in Figure 5 and Table I. As shown, FAU is constructed by three convolutional layers (i.e., Conv3_1, Conv3_2 and Conv3_3) followed by a softmax layer. Symbolically, for each local point patch, let \( W = \{w_1, w_2, w_3, w_4\} \) be the set of learned weights for the four 2D field maps. The FAU can be simply summarised by the following equation:

\[
 f'_{i} = w_i \cdot f_i = FAU(f_i) \cdot f_i \quad (3)
\]

The final aggregated feature is then computed as:

\[
 F^* = Conv(concat[w_1 \cdot f_1; w_2 \cdot f_2; w_3 \cdot f_3; w_4 \cdot f_4]) \quad (4)
\]

FAU assigns different importance to each field 2D maps, allowing the network to pay attention to the salient feature maps in every field. In this way, our FAU can accordingly generate better representations of the shape descriptors which can result in improved classification performance in various complex cases.

Using RRCNet with FAU, each point will be assigned with a probability value \((0 \sim 1)\) which indicates it is belonging to rivet or not. We then take all the points whose probability values are higher than 0.5 as the rivet points. An Euclidean clustering algorithm is then performed to produce a set of clusters, where each cluster is a set of points that are considered to contain one rivet. We finally finish this step by eliminating those clusters that contains much less points.

C. Network training.

We use the manually labeled rivet and non-rivet points from the dataset we collected as supervisory signal. Since our architecture is a binary classification network, binary cross-entropy is adopted as our loss function. Nevertheless, compared to the number of non-rivet points, the number of rivet ones is much small. Thus, a weighted cross entropy loss which emphasizes the loss on rivet points more is employed:

\[
 H = \frac{1}{N} \sum_{i=1}^{N} w_r \cdot l_i \cdot \log (p(l_i)) + (1 - l_i) \cdot \log (1 - p(l_i)) \quad (5)
\]

where \( l_i, i = 1, 2, \ldots, N \) is the label for point \( p_i \) (1 for rivet points and 0 for non-rivet points) and \( p(l_i) \) denotes the predicted probability of the point being rivet for all \( N \) points. \( w_r \) weights the cross-entropy terms for rivet points. Specifically, the weight \( w_r \) is determined by the ratio of the number of non-rivet points and the number of rivet points, that is

\[
 w_r = \frac{\sum_{i} [l_i = 0]}{\sum_{i} [l_i = 1]} \quad (6)
\]

In such way, we force our network to focus on the feature learning on rivet points by penalizing more on misclassifications of them.

VI. RIVET FLUSH MEASUREMENT

Based on the extracted rivet regions, we introduce the rivet flush measurement process as illustrated in Figure 6(a). According to the rivet flush measurement model in Figure 6(b), the rivet flush is defined as the distance between the rivet head and the plane of the surrounding non-rivet points. Therefore, we need to first extract the points on the rivet head. Nevertheless, since the rivet head points are so close to the plane of non-rivet points, as shown in Figure 1, it is difficult to exactly fit the plane of rivet head by a one-step method. To address this issue, we propose a hierarchical fitting method to gradually get the exact rivet head points.

Specifically, for a given rivet region, the hierarchical fitting method consists of three steps, as shown in Figure 6(a). First, the rivet contour point set is extracted via a RANSAC based circle fitting method. Further, rivet head points are extracted.

### Table I

<table>
<thead>
<tr>
<th>Modules</th>
<th>Layer Name</th>
<th>Input Size</th>
<th>Parameters</th>
<th>Output Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN1</td>
<td>Conv1_1</td>
<td>56 × 56</td>
<td>3 × 3, 16, stride 1</td>
<td>56 × 56 × 16</td>
</tr>
<tr>
<td></td>
<td>Conv1_2</td>
<td>56 × 56 × 16</td>
<td>3 × 3, 32, stride 2</td>
<td>28 × 28 × 32</td>
</tr>
<tr>
<td></td>
<td>Pool1</td>
<td>28 × 28 × 32</td>
<td>2 × 2, stride 2</td>
<td>14 × 14 × 32</td>
</tr>
<tr>
<td></td>
<td>Conv2_1</td>
<td>28 × 28 × 32</td>
<td>3 × 3, 64, stride 1</td>
<td>14 × 14 × 64</td>
</tr>
<tr>
<td></td>
<td>Conv2_2</td>
<td>14 × 14 × 64</td>
<td>3 × 3, 128, stride 1</td>
<td>14 × 14 × 128</td>
</tr>
<tr>
<td></td>
<td>Pool2</td>
<td>14 × 14 × 128</td>
<td>2 × 2, stride 2</td>
<td>7 × 7 × 128</td>
</tr>
<tr>
<td>FAU</td>
<td>Conv3_1</td>
<td>7 × 7 × 128</td>
<td>3 × 3, 64, stride 1</td>
<td>7 × 7 × 64</td>
</tr>
<tr>
<td></td>
<td>Conv3_2</td>
<td>7 × 7 × 64</td>
<td>3 × 3, 32, stride 1</td>
<td>7 × 7 × 32</td>
</tr>
<tr>
<td></td>
<td>Conv3_3</td>
<td>7 × 7 × 32</td>
<td>3 × 3, 2, stride 1</td>
<td>7 × 7 × 2</td>
</tr>
<tr>
<td>CNN2</td>
<td>Conv4_1</td>
<td>7 × 7 × 512</td>
<td>3 × 3, 128, stride 1</td>
<td>7 × 7 × 128</td>
</tr>
</tbody>
</table>

*Fig. 5. Architecture of the proposed Field Attention Unit (FAU). Conv3 contains three convolutional layers, each of which is followed by a batch normalization and a ReLU activation function. The parameters of FAU are given in Table I.*
Step 1: RANSAC based circle fitting

Step 2: Rivet head Extraction & finetune

Step 3: Flush computation

(a) Pipeline of measurement

(b) Rivet flush model

Extracted rivet head

Surrounding surface

\[ d_{\text{min}} \quad d_{\text{max}} \]

Fig. 6. (a) Pipeline of the novel rivet flush measurement, which consists of three steps. The most important step is the rivet head extraction & finetune achieved by the proposed hierarchical fitting method. (b) Rivet flush measurement model in our paper.

Step 1: Rivet contour extraction. For a given extracted point set containing one rivet, we first need to find the exact location of the rivet. In fact, the rivet contour can be regarded as the circle pattern in 3D space. As shown in Figure 6(a), the circle pattern is so obvious in the rivet region that we decide to adopt a RANSAC-based circle fitting algorithm for rivet contour detection. Specifically, we first get the mean curvature field of the input point cloud. Further, every point is checked for the curvature value. If the curvature is less than threshold value \( C_t \), then the point is discarded as plane point. \( C_t \) is set to be 0.1 in our experiments. In this way, the remaining points can be seen as the points of rivet contour with few outliers. Finally, we adopt a RANSAC-based 3D circle fitting method to extract the parameters \( (x, y, z, r, n_x, n_y, n_z) \) of the rivet contour circle. \((x, y, z)\) are the coordinates of the circle’s center. \(r\) is the circle’s radius and \((n_x, n_y, n_z)\) are the coordinates of the normal’s direction.

Step 2: Rivet head extraction. According to the rivet flush calculation model in Figure 6(b), we need to extract the rivet head points before the final calculation. With 3D circle parameters \( (x, y, z, r, n_x, n_y, n_z) \) obtained in step 1, a 3D cylinder can be accordingly determined. Rivet head points can then be easily obtained by checking each point if it is inside the cylinder. However, there may exist some outliers in rivet head points, as shown in Figure 7(a). If we use these points to compute the distances to the skin plane, the min distance \( d_{\text{min}} \) would be incorrect. Therefore, we propose to perform a further RANSAC-based plane fitting to obtain the real rivet head points. Since the outliers are so close to the real rivet head that we need to set a much small value for the width of error band. In our experiments, we set it to 0.01. An example of the improved method is given in Figure 7(b).

Step 3: Rivet flush measurement. Through the above two steps, the rivet region is now divided into three parts: rivet head, rivet contour and non-rivet points. Since the aircraft skin surface can be seen as the plane in the local rivet region, we perform the RANSAC-based plane fitting algorithm on the non-rivet points to recover the real surface of the skin. With the fitted plane, we then compute distances between points on the rivet head and the plane. The rivet flush values, i.e., min and max distances \( (d_{\text{min}}, d_{\text{max}}) \), can be easily obtained.
Finally, the distance between the rivet head center and the plane (i.e., rivet flush) is computed as:

$$\text{Flush} = \frac{d_{\text{min}} + d_{\text{max}}}{2} \tag{7}$$

This is a simple yet efficient way to compute the flush since the exact center of the rivet head is difficult to be determined.

VII. RESULTS AND DISCUSSIONS

A. Implementation

For the RRCNet part, we implement the network with Pytorch and train them on a NVIDIA GTX 1080 GPU. The network is trained with the Adam optimizer [42] with an initial learning rate of 0.001 decreased by 0.9 at every 5 epoches. We set beta coefficients to \((0.9, 0.999)\) and batch normalization momentum to 0.5 decreased by 0.5 every 5 epoches. The batch size is set to 32. The curves of accuracy and loss for both training and testing are given in Figure 8. For the data generation and rivet flush measurement, we implement the algorithms with PCL [43] in C++.

B. Dataset

To train our network and evaluate our framework in rivet flush measurement, we use the designed auto-scanning system to collect several point clouds as our dataset for training and testing, as shown in Figure 9. As shown in Figure 9 (b), the left two scanned point clouds in the first column are collected from two standard parts that was specially designed and machined for evaluation of rivet flush measurement. The right two point clouds are captured from the real aircraft skin surfaces. We then annotate these point clouds by manually segmenting the rivet points.

C. Evaluation on Rivet Region Extraction

Evaluation Metrics. Loizou et al. [44] introduced several metrics to evaluate the performance on their boundary point classification from 3D point cloud. Inspired by their work, we adopt precision, recall and IoU to act as the evaluation metrics in this paper.

- **Precision**: It is defined as the percentage of the real predicted rivet points (i.e., True Positives, \(TP\)) in all the predicted rivet points (\(TP + FP\)), that is, \(TP/(TP + FP)\). \(FP\) represents the non-rivet points that are misclassified as rivet points. A higher precision indicates a better performance of the model.

- **Recall**: It is defined as the percentage of the real predicted rivet points (\(TP\)) in all the real rivet points (\(TP + FN\)) in the test data, which is \(TP/(TP + FN)\). \(FN\) is the rivet points which are misclassified as non-rivet points. Recall measures the ability of a model to find all the rivet points within a test set.

- **Rivet IoU (rIoU)**: It is the Intersection over Union (IoU) that measures the area of overlap between the predicted rivet points and the ground truth rivet points divided by the area of union between the above two sets.

- **F1-score**: It is the harmonic mean of precision and recall, that is, \(2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}\).

Comparison. To verify the effectiveness of the proposed components in RRCNet, we compare the performances of our network under different combination of all the three modules (i.e., Multi-field input, Field Attention Unit and Weighted-loss) in Table II. The baseline model takes merely the height map as input, and uses the normal loss function without the extra weight on rivet points. It does not contain field attention unit either. Moreover, we experiment other five kinds of network architecture denoted as RRCNet-A(-D), which contain one or two components, as shown in Table II. RRCNet is the proposed network with all three components embedded in. To make the comparison fair, all comparing networks are trained keeping all hyper-parameters the same.

As shown in the table, we gain 8.9% and 8.3% improvements in terms of Precision from Multi-field input (RRCNet-A) and Weighted-loss (RRCNet-B), respectively. Note that there is not results on RRCNet with FAU individually, since it cannot inserted into the baseline model without multi-field input. By combining multi-field and field-attention unit (RRCNet-C), we can improve the Precision by 9.6%, achieving 92.8%. Weighted-loss brings a further improvement.
TABLE II

<table>
<thead>
<tr>
<th>Model</th>
<th>Multi-field input</th>
<th>Field-attention unit</th>
<th>Weighted-loss</th>
<th>Metrics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>rIoU</td>
</tr>
<tr>
<td>Baseline</td>
<td></td>
<td></td>
<td></td>
<td>80.8</td>
</tr>
<tr>
<td>RRCNet-A</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>91.0</td>
</tr>
<tr>
<td>RRCNet-B</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>85.6</td>
</tr>
<tr>
<td>RRCNet-C</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>90.4</td>
</tr>
<tr>
<td>RRCNet-D</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>91.2</td>
</tr>
<tr>
<td>RRCNet</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>93.2</td>
</tr>
</tbody>
</table>

The higher the better. rIoU: rivet IoU - %, P: Precision - %, R: Recall - %, F: F1-score - %.

Fig. 10. Examples of the predicted results of the proposed networks with different combinations of components. As shown, RRCNet with all the components achieves the best performance in visual, which demonstrates the effect of these components.

Table III

<table>
<thead>
<tr>
<th></th>
<th>RRCNet-A</th>
<th>RRCNet-B</th>
<th>RRCNet-C</th>
<th>RRCNet-D</th>
<th>RRCNet</th>
</tr>
</thead>
<tbody>
<tr>
<td>First row</td>
<td>93.2</td>
<td>93.4</td>
<td>93.4</td>
<td>93.3</td>
<td>93.8</td>
</tr>
<tr>
<td>Second row</td>
<td>63.9</td>
<td>71.2</td>
<td>61.7</td>
<td>80.4</td>
<td>89.3</td>
</tr>
<tr>
<td>Third row</td>
<td>89.1</td>
<td>88.4</td>
<td>86.3</td>
<td>88.3</td>
<td>91.8</td>
</tr>
</tbody>
</table>

Table II shows the effects of the proposed methods (i.e., multi-field input, field-attention unit, and weighted-loss) on the test dataset. The higher the better. The table includes metrics such as rIoU, precision, recall, and F1-score.

The above experiments demonstrate that all the three proposed components contribute to the improvement in performance of RRCNet.

Table III presents statistical results of examples in Figure 10 in terms of F1-score. The table includes results for different configurations of RRCNet:A, B, C, D, and RRCNet.

Figure 10 gives several qualitative results on different networks. As can be seen, RRCNet with all the three proposed components outputs the most similar results to the ground truth in all cases, while RRCNet-A(-D) either miss several rivet points or incorrectly classify non-rivet points as rivet points. In the first row, the input point cloud is in good condition where point density and height difference are obvious. Hence, all the methods successfully extract the rivet region. For the second row, RRCNet-A(-D) miss points near the rivet contour. The reason could be that there is no density difference between rivet and non-rivet region. That leads to the fact that density field map is useless. Thus, taking the density field map as input could bring in noise information in this case. However, our RRCNet can adaptively assign weights to different input fields according to the information of field map itself. In this case, our RRCNet can still perform well by reducing the importance of density field map. Input point cloud in the last row contains heavy noise around the rivet region. RRCNet-A(-D) incorrectly classify several non-rivet points as rivet points, while our RRCNet is not affected by these noise and achieves the best result with reference to the ground truth. We further show the quantitative comparison in the metric of F1-score for the examples in Figure 10. As shown in Table V, RRCNet with the complete configuration still outperforms the comparing RRCNet with different configurations, which is consistent with the visual performance in Figure 10.
TABLE IV

<table>
<thead>
<tr>
<th>Rivet</th>
<th>GT</th>
<th>Measured value</th>
<th>Mean error</th>
<th>Standard deviation</th>
<th>Maximum error</th>
</tr>
</thead>
<tbody>
<tr>
<td>#1</td>
<td>0.110</td>
<td>0.112</td>
<td>0.002</td>
<td>0.007</td>
<td>0.015</td>
</tr>
<tr>
<td>#2</td>
<td>0.131</td>
<td>0.132</td>
<td>0.001</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>#3</td>
<td>0.149</td>
<td>0.148</td>
<td>-0.001</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>#4</td>
<td>0.176</td>
<td>0.178</td>
<td>0.002</td>
<td>0.003</td>
<td>0.005</td>
</tr>
<tr>
<td>#5</td>
<td>0.141</td>
<td>0.141</td>
<td>0.000</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>#6</td>
<td>0.149</td>
<td>0.152</td>
<td>0.003</td>
<td>0.003</td>
<td>0.006</td>
</tr>
<tr>
<td>#7</td>
<td>0.076</td>
<td>0.076</td>
<td>-0.001</td>
<td>0.002</td>
<td>0.003</td>
</tr>
<tr>
<td>#8</td>
<td>0.136</td>
<td>0.136</td>
<td>0.000</td>
<td>0.003</td>
<td>0.004</td>
</tr>
<tr>
<td>#9</td>
<td>0.108</td>
<td>0.109</td>
<td>0.000</td>
<td>0.003</td>
<td>0.006</td>
</tr>
</tbody>
</table>

Average - - 0.001 0.003 0.006

D. Evaluation on Rivet Flush Measurement

Accuracy Analysis. The goal of this paper is to perform rivet flush measurement. In this section, we propose to evaluate the accuracy of the proposed flush measurement method using the point cloud of the standard part in Figure 11. The flush ground truth for these rivets are measured by a 3D coordinate measuring machine. There are nine rivets in the input point cloud here. To reduce the influence caused by randomness, we calculate the flush values ten times, and take the average over ten measurements as the final flush values. Standard deviations and maximum errors are also given to evaluate the robustness of the algorithm.

The quantitative results of nine rivets are given in Table IV. As can be seen in Table IV, the statistical mean errors of flush measurement are between $-0.001\text{mm}$ and $0.003\text{mm}$, which demonstrates that our method has good accuracy. For the standard deviations, all the measurements are less than $0.003\text{mm}$, except for the #1 rivet, whose standard deviation is $0.007\text{mm}$.

Measurement of Real Parts. We further verify the effectiveness of the proposed rivet flush measurement algorithm on the real aircraft skin surface in Figure 13. We show the visual result and the quantitative results in and Table V. As seen, the standard deviation of $d_{\text{max}}$ and $d_{\text{min}}$ are between...
0.003 and 0.005 for the tested ten rivets on real aircraft skin. The maximum deviations are not exceeding 0.008. Figure 12 gives the details of 10-times measurements results of $d_{\text{max}}$ over the ten rivets. As shown, fluctuations over 10-times measurements are much small, which demonstrate our rivet flush measurement method is stable.

E. Limitation

The proposed framework is expected to behave well with different rivet situations. However, there is still one limitation that has to be discussed. That is, it may fail when the rivet fit perfectly with the skin surface, which means the rivet has no height difference with the surrounding surface. In that case, the rivet has no structure difference in 3D shape. Hence it cannot be recognized by the classification network which takes 3D point cloud as input.

VIII. CONCLUSIONS AND FUTURE WORK

In this paper, we present an automatical framework for the detection and measurement of rivets in the aircraft skin. An auto-scanning system is first designed to collect 3D point clouds of rivet region on aircraft skin surfaces. We then introduce the Rivet Region Classification Network (RRCNet) to automatically extract rivet regions in the scanned point clouds. Furthermore, a Field Attention Unit (FAU) is presented to enhance the representation of shape descriptors in different fields via assigning learned weights to the corresponding fields. The experimental results demonstrate the effectiveness of the proposed RRCNet under different types of rivet scanning data. FAU is also proven to be efficient through the ablation study. Finally, a pipeline with a set of point cloud processing techniques is given to generate the results of flush measurement. Experiments demonstrate that the presented rivet flush measurement framework is of high accuracy and works completely automatically. For now, the scanning path is predetermined. In the future, we plan to design algorithms to automatically generate scanning path according to the shape of the target object.

REFERENCES


Dawei Li is currently working towards the Ph.D. degree at Nanjing University of Aeronautics and Astronautics (NUAA), China. He received his Bachelor degree in Mechanical Design and Manufacturing and Automation from Anhui University of Technology (AHUT) in 2016. His research interests include image processing and machine learning.

Yuan Zhang is currently working towards the Ph.D. degree at Nanjing University of Aeronautics and Astronautics (NUAA), China. He received his Master degree in Computer-Aided Design from NUAA in 2017. His research interests include 3D measurement and Inspection.

Jun Wang is currently a professor at Nanjing University of Aeronautics and Astronautics (NUAA), China. He received his Bachelor and PhD degrees in Computer-Aided Design from NUAA in 2002 and 2007 respectively. From 2008 to 2010, he conducted research as a postdoctoral scholar at the University of California and the University of Wisconsin. From 2010 to 2013, he worked as a senior research engineer at Leica Geosystems, USA. In 2013, he paid an academic visit to the Department of Mathematics at Harvard University. His research interests include geometry processing and geometric modeling.