



# Automated quality control inspection of geometric tip defects in medical needle manufacturing

Xiaoying Wang<sup>1,4</sup> · Casey Jowers<sup>2,3,4</sup> · Maciej Mazur<sup>1</sup> · Alexander Buddery<sup>3,4</sup> · Damon Kent<sup>2,4</sup> · Alireza Bab-Hadiashar<sup>1</sup> · Mark Easton<sup>1,4</sup>

Received: 3 November 2021 / Accepted: 3 October 2022 / Published online: 28 October 2022  
© The Author(s) 2022

## Abstract

The manufacture of In Vitro Fertilization (IVF) needles is subject to the most stringent quality demands. This makes automated inspection challenging due to difficulty in reliably classifying conforming and non-conforming (defective) products due to factors including multidimensional variation of their tip geometry and the lack of an explicit quality standard. In addition, developing an IVF needle image dataset, which broadly contains the visual characteristics of qualified and defective products, is difficult without commissioning large and costly production runs. The most important original contribution of this work is a new solution to investigate and quantify the uncertainty in the quality standard of IVF needles by integrating inter-disciplinary techniques. This work utilizes a low-cost, virtual dataset of synthetic images, generated by the automated photo-realistic rendering of a three-dimensional (3D) parametric model to simulate manufacturing variation. Then, the unknown numerical (critical) quality thresholds are obtained by estimating the relationship between quality response and measurement predictors using an Ordinal Logistic Regression (OLR) algorithm on the synthetic images. The fitted models exhibited increased overall predictive accuracy of up to 11.02% than the machine learning models (available in MATLAB) and could provide objective guidance on classifying specific quality aspects of a product.

**Keywords** Automatic quality inspection · Manufacturing variation · Synthetic image generation · Uncertainty quantification · Ordinal logistic regression

## 1 Introduction

The advanced manufacturing of medical In Vitro Fertilization (IVF) needles, even with a low manufacturing defect rate, requires an entirely manual inspection of every unit to remove all defects from production. Human visual inspection of production part quality is critically important to the

manufacturing of safety-critical medical devices to maintain the most stringent quality standard. However, the manual inspection procedures are often time-consuming and costly compared to the products, and have inherent inconsistency due to person to person variation [1, 2]. With the emergence of Industry 4.0 and Smart Factories leading to the increasing convergence of digital and physical worlds, the quality inspection procedure of IVF needles is also seeking automation development that can efficiently boost the manufacturing speed and quality [1, 3, 4]. However, many types of defects have not been characterized in modern manufacturing [1, 2], including the manufacturing of IVF needles. Thus, the bottleneck of this transformation process is the lack of a quantitative quality standard to explicitly guide the discrimination of qualified and defective products by machine.

Automated digital image-based techniques have been extensively applied in the food industry and the manufacturing of mass products due to the fast, cost-efficient, and objective quality inspection capability [5, 6]. For example,

✉ Xiaoying Wang  
xiaoying.wang@rmit.edu.au

<sup>1</sup> School of Engineering, RMIT University, 124 La Trobe St, Melbourne 3000, VIC, Australia

<sup>2</sup> School of Science, Technology and Engineering, University of the Sunshine Coast, Maroochydore BC 4558, QLD, Australia

<sup>3</sup> Cook Medical, 95 Brandl St, Eight Mile Plains 4113, QLD, Australia

<sup>4</sup> ARC Research Hub for Advanced Manufacturing of Medical Devices, The University of Queensland, Brisbane 4072, QLD, Australia

image processing and computer vision techniques have been used in the quality inspection of wheat varieties, corn germplasm, apple shape, nuts, meat, fish and pizza etc. [7–13]. Machine vision assisted quality inspection techniques have also been widely adopted in mass manufacturing of products. An industrial vision system was proposed using optical, photoelectric sensors and a web camera to automatically determine the number of mayonnaise jars on a cardboard tray and whether they had lids or not [2]. Then, a 3D machine vision system was developed for a multilayer perceptron neural network to provide pattern classification of ten possible defects presented in percussion caps [14]. Template localization, edge extraction and distance measurement were utilized in [15] in determining encapsulating quality of polyethylene terephthalate (PET) bottles. Other examples of mass product manufacturing that have adopted vision-based automation inspection techniques include pill packages, rail surfaces, laser welding surfaces, and Printed Circuit Boards (PCB) [16–19].

Compared with the development of automatic vision inspection techniques in food industry and mass product manufacturing, the application of these techniques to IVF needle tips is very rare in the literature due to the challenge caused by metallic reflections on small sized areas [14]. The medical IVF needle manufacturing industry deals with components that feature complex geometries with small feature sizes, requiring robust inspection systems to ensure stringent quality demands; a task which can be challenging for existing visual inspection systems [4, 14]. Safety lancets have a similar tip size to the studied IVF needles. A vision-based quality inspection system was developed for safety lancets [20]. However, the essential inspection parameters, tip size and needle shift value were predefined by an administrator. The vision system presented by [4] combined pattern matching and measurement tools provided by the vision software package Checkpoint to measure shapes and inspect medical syringe assemblies. The recently published vision system in [21] incorporated a photoelectric detector to collect laser signals reflected by knitting needles and used the ratio of adjacent signal peak-to-peak distances to detect fracture and bending of knitting needles. The examined defects occurred on the bodies of knitting needles. This size defect on the knitting needle bodies was clearly noticeable even with the naked eye, thus it posed no problem for image acquisition compared with the defects on IVF needle tips.

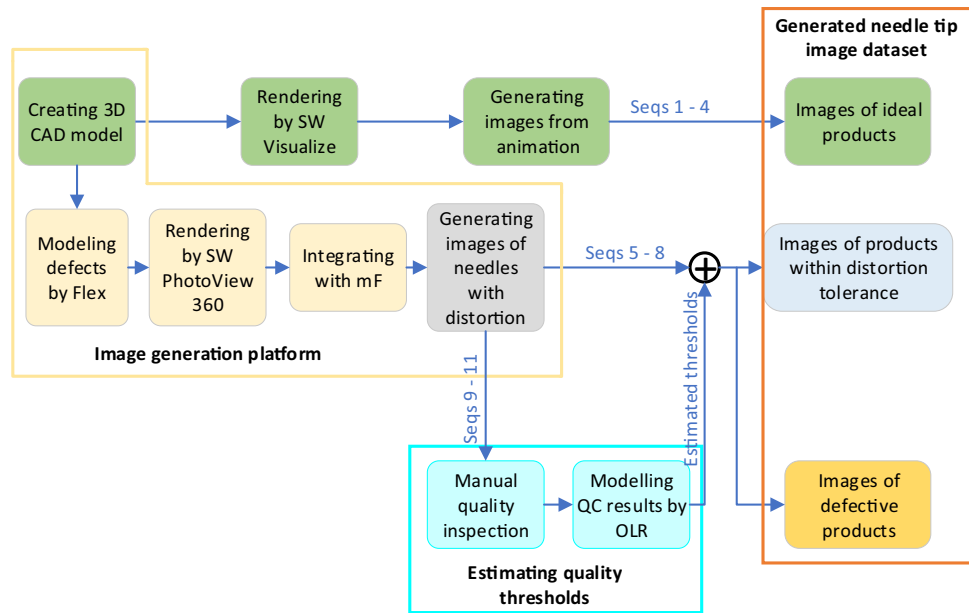
However, none of the work presented above, no matter whether it is in the food industry or modern manufacturing of mass products and small sized medical needles, standardized specification of conforming and non-conforming (defective) products [14, 20]. The critical threshold values were either estimated experimentally from images [2], or given by operators by the distance measurement procedure [15, 20]. To the best knowledge of the authors, the research

on defect/quality standardization of IVF needle tips is also still missing from the literature so far.

Another challenge in studying the defect/quality standard of IVF needle tips from their visual features is the lack of image datasets [22]. Building image datasets requires documenting and analysing the output of manufacturing production runs to characterize the nature of defective products. However, in mature IVF needle manufacturing processes the defect rate is low [23]. Thus, in order to acquire sufficient data for robust defect characterization, it requires a significant period and large production runs to construct a representative dataset covering all types of defects.

A potential way to address the issues of costly image dataset acquisition in visual manufacturing quality inspection, is the use of low-cost synthesized virtual product images in place of comparatively expensive real-life product data. This can be achieved by generating virtual images from parametric 3D Computer-Aided Design (CAD) models which are typically developed during the design phase of the product. Previous research on digital image generation has studied how rendered images can be made to be more photo-realistic [24, 25]. However, generating a sufficiently large dataset of virtual images, which contain photo-realistically rendered part geometry variants of defective parts, remains a labour-intensive task, while automating and standardizing this process is largely an unsolved problem.

This work presents a novel method to quickly and accurately estimate quantitative quality standards of IVF needle tips with minimum resources. The original contributions of this work include the proposed procedure to investigate the quality uncertainty, the identified explicit relationship of needle quality and its geometry information, an innovative approach to adopt the Ordinal Logistic Regression (OLR) algorithm, and a constructed image dataset of IVF needles. First and most importantly, the computational relationship between the predictive product quality and its geometric variables is represented by rigorous equations. During this process, the OLR technique was adopted to [26] model the grading of the generated image dataset into quality categories ‘pass’, ‘unsure’ and ‘fail’. By utilizing the value order characteristic of the algorithm, two mathematical equations were obtained, whose equality demonstrated the result validity. Finally, a large scale, synthesized, photo-realistic image dataset of both specification-conforming and defective products was generated from parametric 3D CAD models. Here, the virtual images were rendered to closely represent the real-life product geometry, material appearance, lighting, and environmental conditions. The significance of this work is twofold. First, advanced IVF needle manufacturing will benefit from the improved inspection speed of machine-assisted visual examination, along with the detailed quantitative insight, and standardization of quality thresholding.



**Fig. 1** Overall workflow, consisting of two main data flows: the top green flow outlines the image generation of ideal products (image sequences 1–4, details of image sequences as seen in Table 1), and the steps in light yellow box indicate the image creation of products with geometric distortion (image sequences 5–11). Based on the extent of the geometric distortion, image sequences 5–11 should be divided into

images of products within/outside (i.e. defective products) the distortion tolerance. Thus, the blue sub-module was built to estimate the quantitative quality thresholds using image sequences 9–11 by modeling their Quality Control (QC) human inspection results, and the obtained quality thresholds were taken as feedback to guide the automatic categorization of image sequences 5–8

On top of that, the constructed IVF needle image dataset demonstrates the capability of the proposed procedure as an economic and quick alternative to construct an image dataset of geometry-specified manufacturing products. This provides the foundation, particularly for advanced manufacturing (with low defect rate), to explore the potential of artificial intelligence technologies that ubiquitously rely on big datasets in image-based automatic quality inspection. The constructed synthetic image dataset and the rendering models from this work are made freely available online<sup>1</sup>.

## 2 Method

The virtual image database construction consists of two main parts: the CAD model construction, including photo-realistic rendering using SolidWorks (SW) 2019–2020 Education Edition, and the automated image generation process. Depending whether the product has or has not geometric distortion, the automated image generation process differs. As shown in Fig. 1, images of ideal products (the top green flow) and images of products with geometric distortion (the middle flow) were generated by creating animation and by integrating with modeFRONTIER (mF, a platform for

process automation and multiple-objective optimization in engineering design), respectively.

The image generation platform, as shown by the light yellow box in Fig. 1 and detailed in Sect. 2.1, was constructed to generate images of products with geometric distortion. After creating the images of products with geometric variants, the statistical model-building technique OLR [26], specifically used for ordered categorical data, is applied to estimate the relationship between the needle tip quality and tip features based on a sub-group of the synthesized digital images (sequences 9–11 in Table 1), as shown by the blue module (estimating quality thresholds) in Fig. 1. The estimated numerical quality thresholds were utilized as guidance to automatically categorize the generated main image group for products with geometric distortions (sequences 5–8 in Table 1).

### 2.1 Construction of image generation platform

The automatic image generation of needle tip variants is achieved by integrating the 3D CAD models created by SW with the multi-objective optimization platform mF via its SW Node. The geometric defective features are first created in 3D CAD models in SW, as shown in Fig. 1. The extent of defects is then controlled by manipulating the values of feature parameters, and product images of variant geometries can be generated for each parameter configuration. Then,

<sup>1</sup> <https://figshare.com/s/c618dc0c6544c3c3e39b>

**Table 1** Overall information of each sub-group for the synthetic dataset. Note: SWV stands for SolidWorks Visualize and PV for PhotoView 360 as photo-realistic rendering tools. Sequences 1–4: images of ideal products. Sequences 5–8: images of needles with tip deformation. Sequences 9–11: images of needles with tip deformation and used for quality threshold estimation

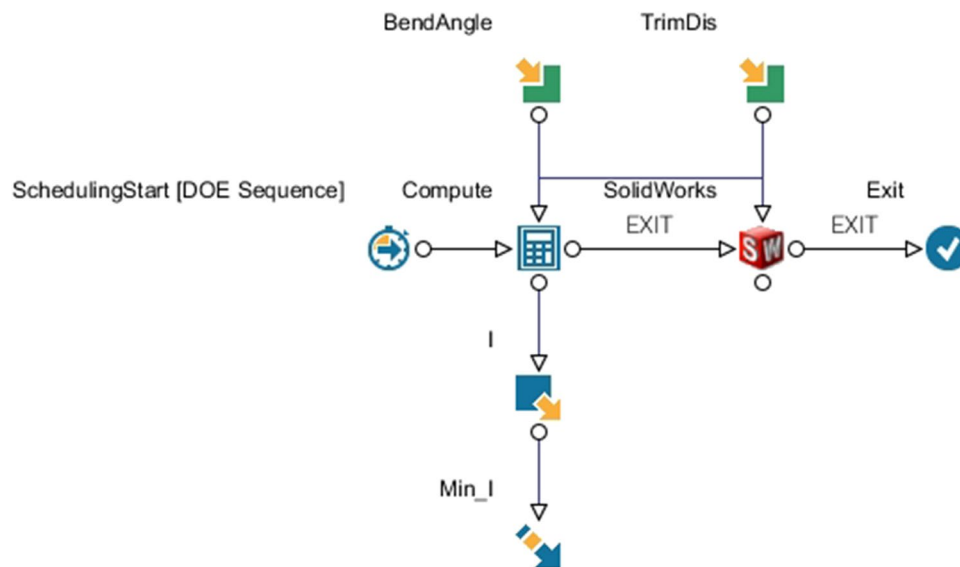
No.	Render tool	$\alpha$ (degree)	$\ell$ (mm)	View point	No. of frames
1	SWV	N/A	N/A	front	1008
2	SWV	N/A	N/A	side	1061
3	SWV	N/A	N/A	side	1230
4	SWV	N/A	N/A	back	1619
5	PV	[2, 60]	[0.1, 1.5]	front	2048
6	PV	[2, 60]	[0.1, 1.5]	left	2048
7	PV	[2, 60]	[0.1, 1.5]	right	2048
8	PV	[2, 60]	[0.1, 1.5]	back	2048
9	PV	[1, 30]	[0.1, 3.0]	side	264
10	PV	[2, 12]	[0.3, 1.8]	side	256
11	PV	[2, 12]	[0.1, 1.8]	side	144

the images are photo-realistically rendered on the 3D CAD models to as closely as possible represent the appearance of materials and environmental conditions. But this only allows one image to be created for a product with fixed geometry and fixed view perspective in SW.

As shown in Fig. 1, integrating 3D CAD models with the multidisciplinary optimization platform, mF, enabled the

automatic generation of large numbers of virtual images of variant product geometry. As shown in Fig. 2, the design of experiment (DOE) sequence created an exploration space, within which the defective feature parameter configuration of the 3D CAD models was distributed. The DOE sequence covered wider values of the geometric defective parameters than anticipated in the real world.

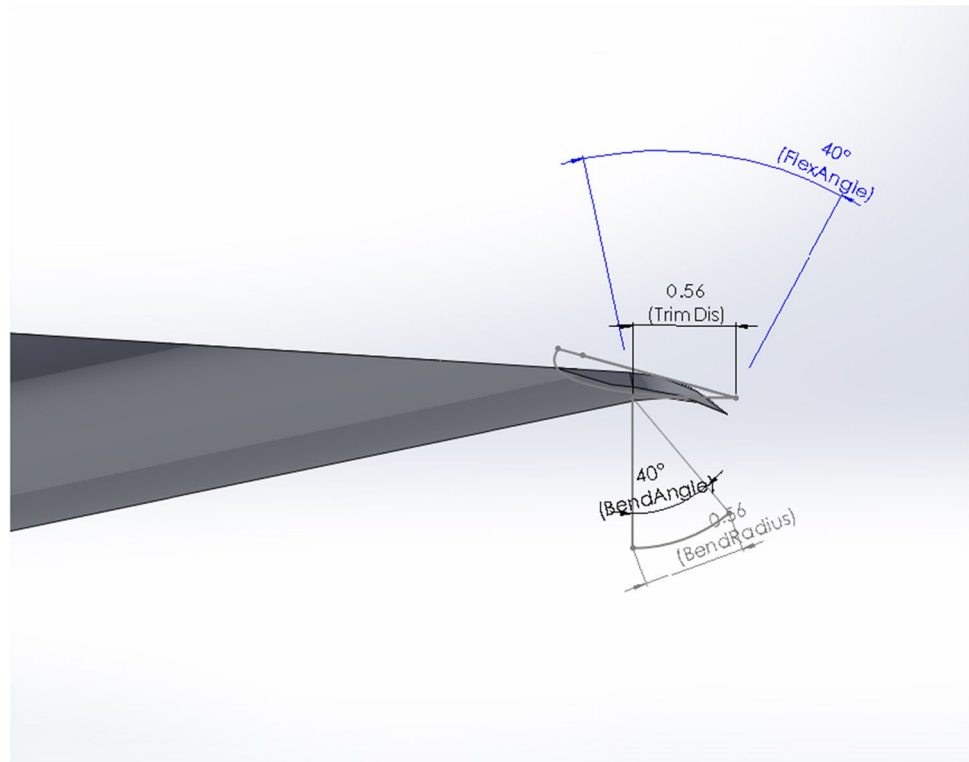
The overall project workflow in mF consists of data flow and logic flow. The stream of data flow (vertical flow) includes input variables, output variables, design objectives and design constraints, as well as the application script with mathematical expressions for calculating the output results. The input parameters as shown in Fig. 2, serving as defective feature parameters in SW, are *BendAngle* ( $\alpha$ ) and *TrimDis* ( $\ell$ ) (indicating the needle defects tip bending degree and tip bending length respectively as shown in Fig. 3). This work focuses on geometric bent defects of needle tips, because the defect identification of bent needles is important for the application in that the IVF needles under study are thick gauge needles so sharpness is critical to ensure the best outcomes for patients. Bend angle  $\alpha$  is a common variable to evaluate the deformation of bevelled needle tips [27, 28]. The other defective feature variable, bending length,  $\ell$ , is identified through quality control inspection in the production line, because the location where a tip bend starts affects the quality control inspection outcome as the closer to the tip end the weaker the bevel strength and the higher the chance for the needle tip to fail. These parameters



**Fig. 2** Overall project workflow in modeFRONTIER (mF), detailing the ‘Integrating with mF’ step in the light yellow box in Fig. 1. The two defective feature parameters *BendAngle* ( $\alpha$ ) and *TrimDis* ( $\ell$ ) in SolidWorks, which define and describe the needle tip bending, are the input variables of mF. Then the output variable  $I$  was computed from bend angle  $\alpha$  and constraint by Eq. 1. All the design configurations

as scheduled in the Design of Experiment (DOE) Sequence were assigned to the two geometric feature parameters *BendAngle* and *TrimDis*. The CAD model geometry automatically updates for each assignment to the feature parameters and rebuilds in the SolidWorks Node within mF. At the same time images of the CAD model related to each design were taken

**Fig. 3** Close examination of tip bend simulated by the Flex feature. FlexAngle and TrimDis are the two default parameters of the Flex feature to directly control the bend degree from the needle main axis and the start of the bend from the tip end. FlexAngle is set to be an equation-driven variable, whose value is equal to the independent reference variable, BendAngle for easy data accessibility and direct data manipulation via SW Equations



are selected from the SW Equations, linking with mF input variables, and then computed by mathematical expressions written in the Calculator Node script in mF. The calculated results are the mF output variables (the variable  $I$  in the example as shown in Fig. 2). Although calculating a quantitative objective (usually used for guiding and optimizing part design within mF) is not one of the aims of this image generation procedure, an optimization objective is preferred when using mF. Thus a pseudo-objective, as shown in Eq. 1, was created in the application script. Because the bend angle created by the Flex feature in SW cannot equal zero,  $\alpha \neq 0$  is the design constraint, which also satisfies the constraint of Eq. 1 whose denominator cannot be zero.

$$\min I = \frac{1}{\alpha^2} \quad (1)$$

After choosing the input variables, their initial values in the 3D CAD models should be updated from SW to ensure that they are within their value ranges for each running of the program. Otherwise, a portion of the input parameter configurations will cause errors in the design space, due to the initial values being out of the range of the input variable.

The logic flow is a sequence of logic events in mF that enables it to solve a defined optimization problem in the

aforementioned data flow. A typical optimization process objective is to find one optimum set of parameters which lead to the desired output. However, in our example, digital images are required from all parameter configuration points, and therefore the logic flow is slightly different. As shown in Fig. 2, the logic flow (horizontal flow) starts from the DOE sequence Scheduler node, then connects with the Calculator node (Compute here) followed by the red SW Node (to integrate the 3D CAD model) and finishes with the logic end node (Exit).

DOE is a systematic method to choose variables for multiple parameters that varies in a series of experiments with the purpose to gain the maximum knowledge of the studied model, while expending the minimum time and costs for computation [29]. The Sobol sequence, incremental space filler and factorial algorithms [30–32] are applied to generate the sequence of different design configurations, which makes up the design space in this work. Each assignment to the parameters of the model is estimated by evaluating the created DOE sequences. The model geometry automatically updates and rebuilds in SW while mF modifies the input values. Because mF enables saving the SW documents with modifications related to each design, images of needle tips were taken from top, left/right side and back perspective views from SW 3D models and stored for every geometry design.

## 2.2 Images of defective products

As shown in Fig. 1 the first step to generate images of defective products is creating a 3D CAD model, including creating a tip bending feature in the CAD model described by tip bend angle and bend length. The images of defective products are then photo-realistically rendered in SW PhotoView 360.

### 2.2.1 Modeling defects

A 3D CAD model of ideal IVF needle geometry is first constructed in SW, with the Bending applied using the Flex feature to simulate the tip bending defect type as shown in Fig. 3. The FlexAngle parameter defines the range between which the object will be affected by the bending feature, which is one essential aspect to control the shape of the resulting bend. The actual bending angle parameter is set to be equation-driven, whose value is equal to a newly created independent reference variable named *BendAngle* ( $\alpha$ ) (initially valued  $40^\circ$  as shown in Fig. 3). The positive or negative sign of the assigned value controls the needle tip bend direction. Similarly, the bend length *TrimDis* ( $\ell$ ), is also set to be equation-driven.

### 2.2.2 Visual realistic rendering in SW

The 3D model was rendered to obtain high quality images (Fig. 4c). Firstly, materials were applied to the 3D model to simulate their appearance. Then PhotoView 360, a SW add-in, was utilized to incorporate realistic material appearance, illumination lighting (primary lighting and multi-directional lighting) and environmental scene with virtual camera perspective view.

The appearance of the inside and outside of the needle cannula (due to the surface finish) is different, so satin finish stainless steel and cast stainless steel were the applied material appearances to the needle body and inside surface, respectively, which closely resemble their actual visual

appearance. Small size cylindrical mapping with  $180^\circ$  orientation was applied to the inside surface of the cannula to simulate the wavy pattern texture and the vertical pattern direction.

## 2.3 Images of qualified products

Synthesized images of qualified products consist of two groups, as shown in Fig. 1. One group contains the virtual images generated using the same process described for the defective products, but with defects insignificant in scale and within pre-determined distortion tolerance limits. The second group of images simulates the ideal manufacturing results, where all geometry features are identical to the 3D CAD models without any geometric distortion.

### 2.3.1 Images of products within distortion tolerance

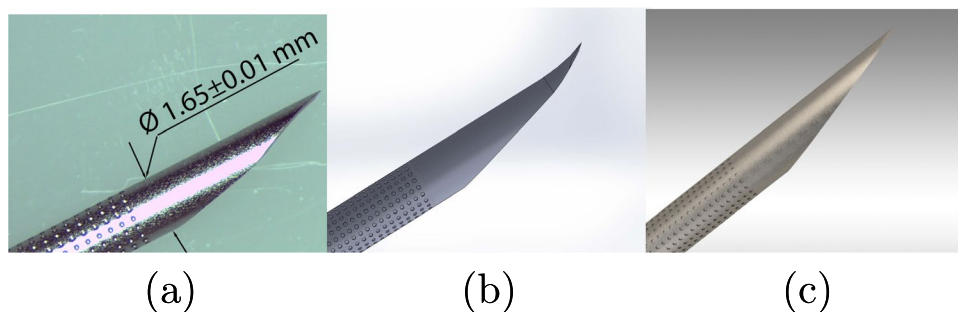
The essential step to collect this group of synthesized images for qualified manufacturing products is to estimate the proper quantitative tolerance thresholds, including the tip bending limits, as indicated by the *BendAngle* ( $\alpha$ ) and *TrimDis* ( $\ell$ ) parameters in Fig. 3. Details on the method used for quality threshold estimation are presented in Sect. 2.4. Many more images were generated across the defined tolerance ranges.

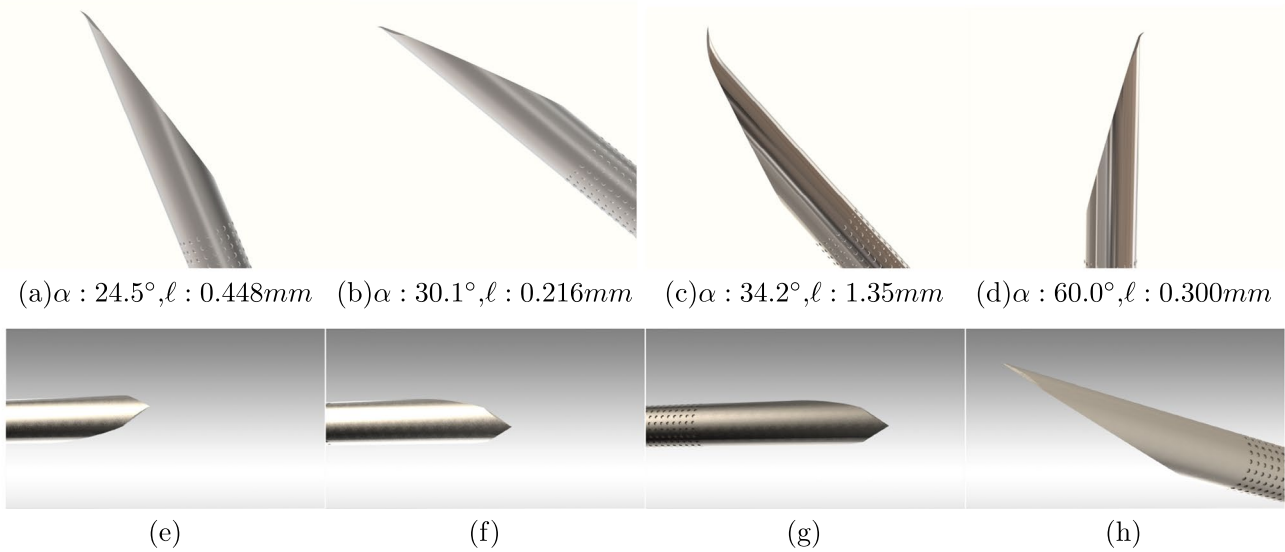
### 2.3.2 Images of ideal products

The image group of ideal products having theoretical geometry was rendered and generated via animations in SW Visualize using the 3D CAD models, as shown in Fig. 1. Without geometric parameters, it is not feasible to create this group of images using mF.

However, SW Visualize provides an approach to automatically generate images for 3D models with fixed geometry by creating animations. In addition to its rendering capabilities with design-oriented features for generating compelling, highly realistic images, SW Visualize has more advanced features and control over material appearances. For example, sandblasted steel instead of satin finish stainless steel

**Fig. 4** **a** Image of a IVF needle taken by an optical microscope; **b** Digitally generated image example with tip bending defect from its 3D CAD model without rendering; **c** Generated image of a qualified needle tip with photo-realistic rendering





**Fig. 5** Examples of the generated virtual images dataset. Images (a)–(d) are examples of needles with variant tip bending selected from image sequences 5–8 of Table 1, while images (e)–(h) of needles

with ideal geometry (selected from image sequences 1–4 of Table 1) and so bend parameters are not applicable

is used to simulate the needle body appearance, which vividly displays the actual appearance of the cannula outside the surface and cutting bevel (as shown in Fig. 5f and g). Additionally, several animations can be stacked on top of each other to cover more comprehensive environmental conditions, including rotating the view perspective, adjusting illumination, trucking or dollying into/out of the point of view, and adjusting the virtual camera locations.

### 2.4 Statistically estimating quality thresholds

The response outcome, needle tip quality, has three ordered categories, containing images of either: qualified needle tips (‘pass’), defective needle tips (‘fail’), or needle tips whose qualities are between the two (‘unsure’); therefore OLR (also called ordered logit model or proportional odds model [33]) was used to model the mathematical relationship between the ordinal quality response and the predictors (defective feature parameters bend angle  $\alpha$  and bend length  $\ell$  of needle tips). OLR is a competitive Statistical Regression (SR) algorithm for model-building in classification and prediction tasks [26]. This technique has been widely used in many industrial applications over the last three-to-four decades [34–37]. Compared to machine learning models, OLR does not require such a large number of sample images. The OLR analysis is conducted in the statistical software Minitab [38].

The batch of 264 images (sequence 9 in Table 1) covering wide ranges of values of bend angle  $\alpha$  and length  $\ell$  are firstly used to estimate the possible threshold ranges. Then another two image batches (sequences 10 and 11 in Table 1) are further generated around the narrowed threshold ranges

to increase the density of the data points near this area. There are 664 images in the three batches, whose orders have been randomized prior to human QC. These images have been manually classified into the three categories (‘Pass,’ ‘Fail,’ and ‘Unsure’) by a subject matter expert. The needle tips in the images are visually examined at a pre-determined scale (replicating that of a 15× Optical Microscope) to determine whether there is any damage to the tip (or at least any damage which is visible at this pre-determined scale). The acceptance criteria used to qualify the needles is based on the manufacturer’s Standard Operating Procedure; however, in this instance, a third category (‘Unsure’) is used to estimate and account for inherent process variability (including differences between the operators or inspection equipment). Consequently, by structuring the discretization of the model to contain three categories (with two critical thresholds), we improve insight and control of the model. In a medical device manufacturing application, these might translate practically to the following classification labels: ‘Definitely Pass’, ‘Definitely Fail’, and ‘Further Human QC Required.’ The QC results of the 664 images are used to determine the threshold ranges according to the bend angle  $\alpha$  and bend length  $\ell$  variables.

#### 2.4.1 Ordinal logistic regression

OLR works with the cumulative response probabilities [34]:

$$\gamma_1 = \pi_1, \gamma_2 = \pi_1 + \pi_2, \dots, \gamma_k \equiv 1 \tag{2}$$

$$\gamma_j = pr(Y \leq j) = \pi_1 + \pi_2 + \dots + \pi_j \tag{3}$$

where  $\pi_1, \dots, \pi_k$  are the category or possible event probabilities, and  $k$  is the number of possible events. There are three possible event categories, so  $k = 3$  here. Note that the categories are ordered,  $Y$  is the event combination up to the  $j^{th}$  event. Thus,  $\gamma_j = pr(Y \leq j)$  is the cumulative probability of up to and including  $j^{th}$  event.

In particular, the logistic link function (as shown in Eq. 4) for cumulative probabilities has been found to have better properties than categorical probabilities for ordered categorical data, including permutation invariance (i.e. the fitted model is invariant under a reversal of category order) and constant difference [26]. Also, the logistic link function has the most natural interpretation of the estimated parameters compared to normal or Gompertz distribution [26, 39].

$$\ln\left(\frac{\gamma_j(\mathbf{X})}{1 - \gamma_j(\mathbf{X})}\right) = \theta_j + \beta^T \mathbf{X}, j = 1, 2, \dots, k - 1 \tag{4}$$

where  $\mathbf{X}$  is the covariate vector, consisting of bend angle  $\alpha$  and bend length  $\ell$  of a needle tip;  $\theta_j$  and vector  $\beta$  are the parameters of the model and determined by datasets. Considering Eq. 3 and applying exponentiation for both sides of Eq. 4, we get Eq. 5:

$$pr(Y \leq j) = \gamma_j = 1 - \frac{1}{1 + e^{\theta_j + \beta^T \mathbf{X}}}, j = 1, 2, \dots, k - 1 \tag{5}$$

According to the order arrangement characteristic, there are two ways to calculate the probability equation of the interested response event (i.e. defective needle tips). If the response is ordered as ‘pass’, ‘unsure’, and ‘fail’:

the ‘pass’ probability is

$$\pi_j = pr(Y \leq j) = \gamma_j = 1 - \frac{1}{1 + e^{\theta_j + \beta^T \mathbf{X}}}, j = 1 \tag{6}$$

and the cumulative probability of ‘pass’ and ‘unsure’ is

$$pr(Y \leq j) = \gamma_j = 1 - \frac{1}{1 + e^{\theta_j + \beta^T \mathbf{X}}}, j = 2 \tag{7}$$

Thus, the ‘fail’ probability is given by

$$1 - pr(Y \leq j) = \frac{1}{1 + e^{\theta_j + \beta^T \mathbf{X}}}, j = 2 \tag{8}$$

The other method of calculating the ‘fail’ probability is implemented in this work by arranging the response order as ‘fail’, ‘unsure’, and ‘pass’ due to the permutation invariance property of the proportional logistic function. Then the ‘fail’ probability is obtained directly from Eq. 6. Here when  $j = 1$  under the reversal of categorical order, the event it represents is ‘fail’ instead of ‘pass’.

This work implements both the models annotated by Eqs. 8 and 6 to estimate the relationship between the

predictors and the quality fail probability. We should obtain the identical fitting results from both models (represented by different mathematical equations), if the above statistical process is correct.

### 2.4.2 Correlation test

To predict the response variable, both models are formulated in terms of their two predictors, bend angle and bend length. It is necessary to examine the severity of their correlation (also known as multicollinearity [33, 40]). The correlation value  $r$ , as shown in Eq. 9 and calculated using Pearson product moment method [38, 41, 42] is 0.157 with 95% confidence interval (0.082, 0.231), and the  $\rho$  value is less than 0.005, which is much less than the required significance level of 0.05. Thus, there is sufficient evidence to say there is not a strong linear relationship between the two predictor variables.

$$r = \frac{\sum(X - \bar{X})(Y - \bar{Y})}{(n - 1)S_X S_Y} \tag{9}$$

where  $\bar{X}$  and  $\bar{Y}$  are the means of the two predictor variables, bend angle and bend length, and  $S_X$  and  $S_Y$  are the standard deviations of the two variables.

## 3 Results and discussion

### 3.1 Generated needle tip image dataset

Table 1 briefly summarizes the overall information of the synthesized needle tip image dataset. A total of 13, 774 images were generated. Sequences 1–4 are images of ideal needles without tip deformation, so both bend angle and bend length are zero. Undoubtedly, these four image sequences belong to the ‘pass’ quality category. While sequences 5–8 accommodate images from the three possible quality categories based on their tip bending values. Thus, the sub-group of 664 frames, sequences 9–11, were taken to estimate the numerical quality threshold of needle tips and the estimated quality threshold is then used to categorize sequences 5–8, also as shown in Fig. 1.

Figure 5a–h show some image samples rendered by PhotoView 360 ((a)–(d)) and SW Visualize ((e)–(h)) from the created dataset. Both groups of images are taken from variant environmental lighting, perspective views, orientations, and distance from the virtual lenses. All the rendered images manifest in photo-realistic appearance compared with the 3D model without rendering as shown in Fig. 4b. It can also be seen that images in group Fig. 5e–h display higher photo-realism with more accurate rendering display and lighting



reflection. The outside surfaces of the cannula using the sandblasted steel appearance, as shown in group Fig. 5e–h, show an improved visual appearance when compared to group Fig. 5a–d.

### 3.2 Statistical analysis on the threshold of needle tip quality

One of the important contributions of this work was estimating the numerical thresholds of two critical defective feature parameters. The estimated thresholds are also directly used in the automated categorization of image sequences into their quality classes to make the presented virtual image generation procedure scalable.

#### 3.2.1 Initial visualization of the quality results

The 664 frames of images (sequences 9–11) used for estimating the numerical threshold of needle tips consist of three sequences and Fig. 6 displays the quality control results of these needle tip images. The 1st batch of 264 images (sequence 9) covers a broad range of the predictor variables  $\alpha$  and  $\ell$ , as shown in Fig. 6a, to investigate the

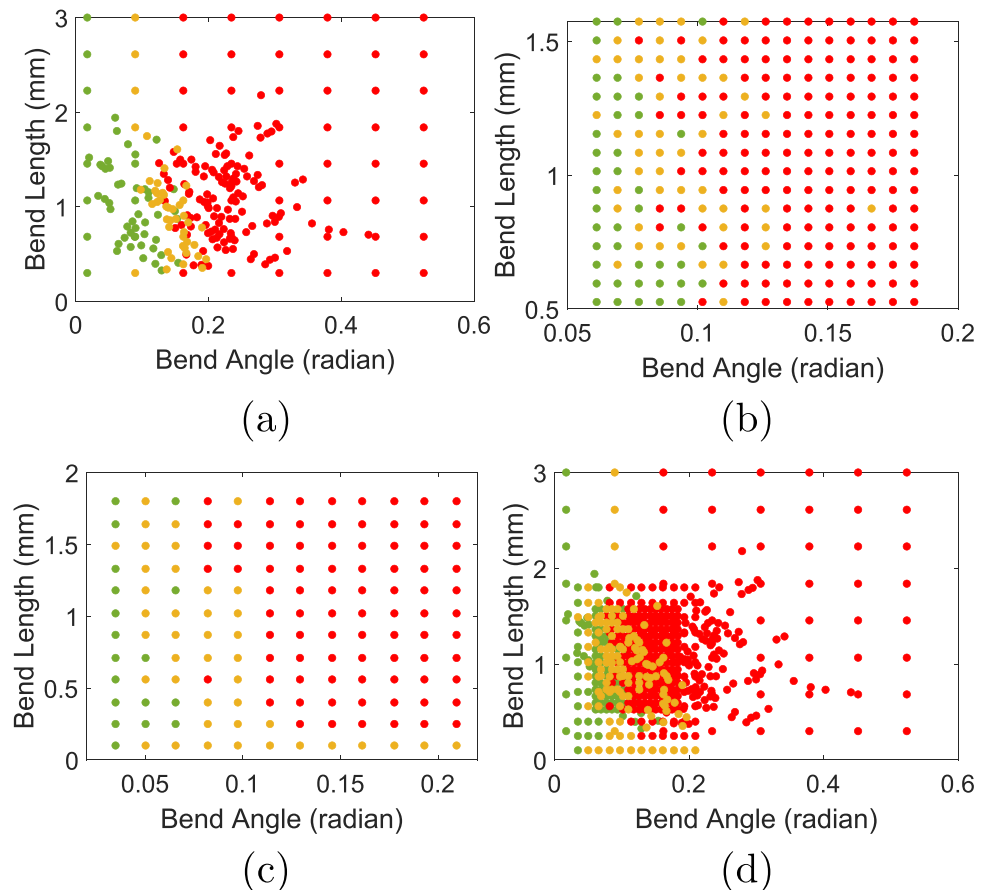
threshold area. The mixture of green, yellow and red dots as shown in Fig. 6a are the identified threshold ranges. To increase the accuracy of the estimated quality threshold, another two batches of images (sequences 10 and 11) were generated around the identified threshold ranges. Figure 6b and c show the quality control results of these two sequences of images and Fig. 6d summarizes the quality control results using all the three sequences of the images.

The visualization of the initial result in Fig. 6 illustrates the consistency of needle tip quality along the predictor variables  $\alpha$  and  $\ell$ . Also, the bend angle  $\alpha$  affects the quality determination of a needle tip more than the bend length  $\ell$ .

#### 3.2.2 Ordinal logistic regression results

Tables 2 and 3 summarize the results of the OLR analysis using Eqs. 8 and 6, including the estimated coefficients, the standard errors of the coefficients, Z-values, P-values, as well as the odds ratios and their 95% confidence intervals. These two tables also provide the results of goodness-of-fit tests to measure the adequacy of the regression models.

**Fig. 6** Visualization of the quality control results using image sequences 9–11 as shown in Table 1 and Fig. 1. Green, yellow and red dots annotate ‘pass’, ‘unsure’ and ‘fail’ quality results, respectively. **a** Quality results covering broad ranges of predictor variables  $\alpha$  and  $\ell$ ; **b** and **c** Quality results for covariate values near the narrowed threshold area; **d** Summarized quality results of the three image sequences on one figure



**Table 2** Results of ordinal logistic regression using Eq. 8

Predictor	Coef	SE(Coef)	Z	P	Odds Ratio	95% CI	Goodness-of-fit tests		
$\theta_1$	-6.76	0.510	-13.3	< 0.0005			Method	Chi-Square	P
$\theta_2$	-4.63	0.447	-10.4	< 0.0005			Pearson	938	1.00
$\alpha$	46.4	3.10	15.0	< 0.0005	$1.48 \times 10^{20}$	$[3.37 \times 10^{17}$ $6.47 \times 10^{22}]$	Deviance	739	1.00
$\ell$	1.22	0.234	5.21	< 0.0005	3.39	[2.14 5.36]			

From Table 2 the estimated parameters for Eq. 8 are  $\theta_1 = -6.76$ ,  $\theta_2 = -4.63$  and  $\beta^T = [46.4, 1.22]$ . Thus, the computed ‘fail’ probability using Eq. 8 is

$$1 - pr(Y \leq j) = \frac{1}{1 + e^{-4.63+46.4*\alpha+1.22*\ell}} \tag{10}$$

where  $j = 2$ , and  $\alpha$  and  $\ell$  stand for bend angle in radians and bend length in mm. Bend angle was transformed to radians to keep the same degree of magnitude/range with bend length (feature scaling) for more accurate model building.

The calculated coefficient of 46.4 for bend angle  $\alpha$  is the estimated change in the logit of the ‘fail’ probability when bend angle changes by one unit, with the covariant bend length  $\ell$  held constant. The estimated positive coefficient of bend angle  $\alpha$  indicates that as the bend angle increases, the ‘fail’ probability also increases. The positive bend length  $\ell$  coefficient indicates its similar effect on the needle quality, although with less severity. These conclusions are consistent with the experiences obtained in practice. The less than 0.0005  $p$ -values for both bend angle ( $z = 15.0, p < 0.0005$ ) and bend length ( $z = 5.21, p < 0.0005$ ) indicate that there is sufficient evidence that both predictors have non-zero values at 0.0005 significance level. Thus, both bend angle  $\alpha$  and bend length  $\ell$  have an effect on determination of needle tip quality.

Goodness-of-fit tests were conducted to assess the adequacy of the estimated model to the data. If the estimated model does not fit the data well, then the classification and prediction results using the model can be misleading. The higher the Chi-Square value the more adequate of the model to the data. The null hypothesis of the Goodness-of-fit test is the model is not a good fit, so a higher  $p$ -value indicates there is insufficient evidence to support the null hypothesis. Thus, the high  $p$ -values (1.000) from

both Chi-Square Pearson and Deviance statistics methods [43–45] illustrate that there is insufficient evidence to claim that the model does not fit the data adequately.

Table 3 shows the regression results using Eq. 6 and the results of goodness-of-fit tests to the regression relationship. The calculated ‘fail’ probability using this equation is

$$\pi_j = 1 - \frac{1}{1 + e^{4.63-46.4*\alpha-1.22*\ell}} \tag{11}$$

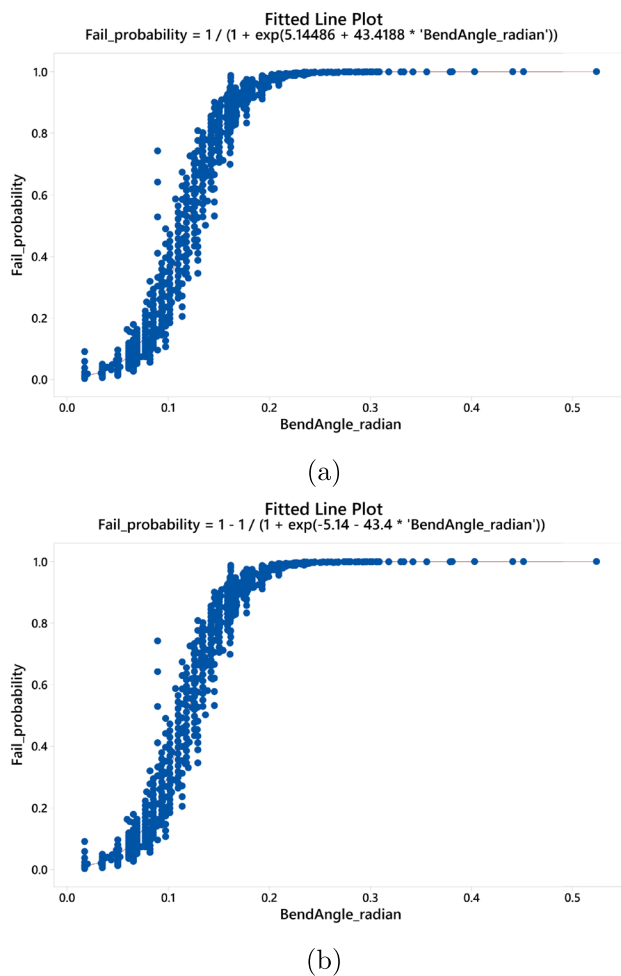
where  $j = 1$ , and  $\alpha$  and  $\ell$  stand for bend angle in radians and bend length in mm. Because this group of estimated parameters is obtained by rearranging the category orders, based on the same data samples, the calculated parameters are the negative values of Table 2. For example, the estimated negative coefficient of bend angle  $\alpha$  and an odds ratio that is smaller than 0.0005 (shown as 0.00 in the table) indicate greater bend angles of needles are associated with a smaller probability of the non-fail quality acceptance criteria, which results in an increased probability to fail the needle quality. Also, the much higher coefficient of bend angle  $\alpha$  than bend length  $\ell$  indicates the greater effect of bend angle than bend length on needle tip quality.

As shown in Sect. 2.4.1, if the statistical process is correct, the two estimated models of Eqs. 10 and 11 should be identical. Thus, the calculated ‘fail’ probabilities from these two equations are shown in Fig. 7a and b for comparison. The models are simplified by dropping the small-valued term of bend length to display the results graphically. It can be seen that the estimated models represent the same relationship between the quality response and the covariates of bend angle  $\alpha$  and bend length  $\ell$ .

We are interested in the ‘fail’ response event of a needle tip quality in the manufacturing process. Thus, Fig. 8 presents the final estimated model of the ‘fail’ probability

**Table 3** Results of ordinal logistic regression using Eq. 6

Predictor	Coef	SE(Coef)	Z	P	Odds Ratio	95% CI	Goodness-of-fit tests		
$\theta_1$	4.63	0.447	10.4	< 0.0005			Method	Chi-Square	P
$\theta_2$	6.76	0.510	13.3	< 0.0005			Pearson	938	1.00
$\alpha$	-46.4	3.10	-15.0	< 0.0005	0.00	[0.00 0.00]	Deviance	739	1.00
$\ell$	-1.22	0.234	-5.21	< 0.0005	0.30	[0.19 0.47]			



**Fig. 7** Fitted lines of the simplified models: **a** from Eq. 11 and **b** from Eq. 10 for the ‘fail’ probability of needle tip quality

of the needle tips based on the values of tip bend angle  $\alpha$  and bend length  $\ell$  in a 2D contour plot. As shown in Fig. 8 the contour is not linear, because for bend length  $\ell > 2$  mm and bend angle  $\alpha > 5$ deg it is more likely to fail (because the defect is more apparent to the eye/brain), but for bend length  $\ell > 2$  mm,  $\alpha < 5$ deg, it is more likely to pass because of the increased cross sectional area of the bevel. When the bend angle and bend length of a needle tip is known (e.g. by measuring the relative image pixels of a needle tip), its quality can be easily predicted or retrieved using the figure. The estimated ‘fail’ probability as shown in Fig. 8 also provides further insight into different confidence acceptance levels to the decision-making process.

### 3.2.3 Comparison with machine learning models

The images in sequences 9–11 in Table 1 that are used to estimate the numerical threshold of needle tip quality have

been also trained and tested using machine learning models. The selected results from the automated classifiers [46–49] in MATLAB with the top four predictive accuracy are tabulated into Table 4. The machine learning models are validated using 5-fold cross validation during training to protect against overfitting based on 80% of the image batches. The predictive accuracy of the fitted models is further examined by the withheld 20% images (132 frames). Due to the small size of the sub-group images, the fitted models (with the highest predictive accuracy in validation data and median value, 87.50% and 80.45% respectively) obtained by machine learning techniques did not outperform the OLR model. The Friedman test [50] was conducted on the classification accuracy of compared algorithms to determine whether any of the differences between the population medians were statistically significant. The results were tabulated in Table 4. The small  $p$ -value 0.003 (less than the usual significance level 0.05) indicates that there is enough evidence to reject the null hypothesis that all treatment effects are zero. Thus, it can be concluded that not all the population medians are equal.

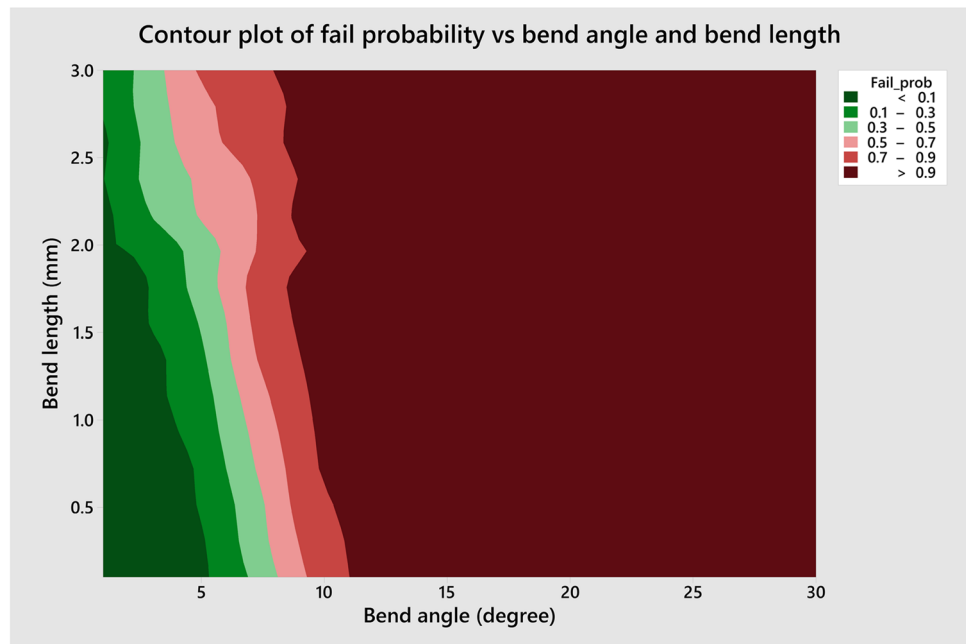
To further examine which algorithm provided a statistically different median value, the post hoc Nemenyi test [51] was run following the Friedman test. As shown in Table 5, the  $p$ -value of the comparison pair, Optimized SVM and the proposed method, is 0.003, which is less than a significance level 0.05. The proposed algorithm exhibited increased median predictive accuracy of 11.02% greater than the optimized SVM. In addition, the sum of ranks of the proposed algorithm, shown in Table 4, is the highest among the methods, illustrating that the proposed algorithm is statistically different and has a higher classification accuracy.

Furthermore, the fitted OLR model provided the numerical threshold for each predictor of needle tips, while the trained machine learning models cannot output such detailed criteria information. By adjusting the numerical thresholds according to the resulted fail probability, the false-positive occurrences can be avoided. Thus, the fitted OLR models satisfy the requirement to numerically estimate the quality thresholds of needle tips.

### 3.3 Manufacturing experiment

This section presents the discrimination of qualified and (bending) defective IVF real-life needles using image processing techniques. We developed a new vision system, consisting of a Blackfly USB3 colour camera mounted on top of a TechSpec PlatinumTL 0.9X telecentric lens. Illumination was especially designed in top and back lights for sufficient projection into the camera and the creation of sharp needle contour. The top CCS dome white light was controlled by a CCS 46W digital power. In addition,

**Fig. 8** Contour plot of ‘fail’ probability vs. bend angle  $\alpha$  and bend length  $\ell$ , estimated from image sequences 9–11 of Table 1. The estimated mathematical relationship of needle tip quality with the measurement predictors, bend angle and bend length, could be utilized as a numerical guide for automatic image-based quality inspection. It was also utilized for automatic (generated) image categorization in this work



a 0.312-inch LED spot white light guided by a TechSpec telecentric illuminator lens was set up underneath needles to create a telecentric backlight. The telecentric lens, telecentric illuminator and dome light were held by universal stands so that their heights could be smoothly adjusted vertically.

A steel base plate was devised and machined to support needles and locate them in a predefined position. When taking images of real-life needles in this dataset, a single needle was placed at a time using one particular groove. Three images from the top, side and back perspective were taken for one needle tip. There were 40 good needle samples and 35 defective samples with tip bending tested, 225 images in total.

Five virtual images with variant bending degrees were selected as template candidates. Tip areas of the side view were

cropped as the final templates. First, images of real-life needles were converted into grey-level images from (Red Green Blue) RGB colour images. Then the image edge features were calculated using horizontal and vertical intensity gradients. The purpose to detect image edges was to eliminate the effect of ambient lighting and only consider geometry shape and texture features. At the same time, five templates were also processed according to the same procedure. The final step calculated the correlation coefficients between the templates and the real-life images using the normalized cross-correlation method [52]. The highest value of the five templates indicated the closest similarity to that template, whose bending parameters determined the quality of the examined real-life needles. 36/40 good samples and 33/35 defective samples were correctly classified, the results of the preliminary experiment showed an overall accuracy of 92%.

**Table 4** Friedman test on the classification accuracy by the proposed and machine learning models, where Q-SVM, MG-SVM, O-SVM and DT stand for Quadratic SVM, Medium Gaussian SVM, Optimized SVM and Decision Tree, respectively

Iteration	Q-SVM [46]	MG-SVM [47]	O-SVM [48]	DT [49]	Proposed
1	79.25	71.70	70.75	74.06	81.60
2	76.17	78.04	77.57	73.83	81.31
3	84.11	84.11	75.23	78.04	88.79
4	76.89	74.06	68.40	69.81	81.60
5	82.55	78.77	75.94	81.60	82.55
Validation	83.33	87.50	68.56	81.82	82.58
Median	80.45	78.98	72.85	75.93	83.87
Sum of ranks	22	19.5	8.0	13	27.5
Friedman test	<i>chi - square</i> = 15.831, <i>df</i> = 4, <i>p</i> -value = 0.0033				

**Table 5** Pairwise comparisons using Nemenyi post hoc test for a two-way balanced complete block design

Method	Q-SVM	MG-SVM	O-SVM	DT
MG-SVM	0.99	-	-	-
O-SVM	0.08	0.22	-	-
DT	0.47	0.76	0.89	-
Proposed	0.85	0.59	0.003	0.06

## 4 Conclusion

This work provides a novel approach to estimate the quantitative standard of medical IVF needles to address the quality uncertainty problem in manufacturing. The absence of standardized quality criteria and the low defective product rate hamper the application of machine-assisted automation techniques to product quality control inspection in the advanced medical manufacturing. Another contribution of this work is the proposed procedure to automatically generate a large scale synthetic image dataset of qualified and defective products based on 3D CAD models, with anti-aliasing (by fully defined defective parameters) and minimal time and resource effort. In addition, the computer-generated images are rendered to a level that is industrially representative and photo-realistic, fully simulating the complex material appearance, illumination, reflection and environmental conditions. Overall, the quantitative quality criteria are estimated and obtained via OLR analysis on the constructed synthetic image dataset. In future work, deep/machine learning algorithms can be primitively evaluated for their capability to classify qualified and defective needle tips using this computer-generated image dataset. If a model performs well on this image dataset, it can be shortlisted in classifying real-life needles. In addition, a machine vision system will be developed to detect more types of defects on real-life IVF needle images.

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1007/s00170-022-10249-5>.

**Acknowledgements** We thank Xan MacAtangay for his helpful advice on applying rendering on 3D CAD models.

**Author contributions** The project was developed by Alexander Buddery and Mark Easton. Xiaoying Wang, Casey Jowers, Maciej Mazur, Alexander Buddery, Alireza Bab-Hadiashar and Mark Easton conceived and designed the experiments. Xiaoying Wang developed software, analysed data and wrote the original draft. Casey Jowers inspected the generated images, labelled their classes and wrote the description of this section. Maciej Mazur guided the simulation activities. All the authors reviewed, revised and edited the manuscript.

**Funding** Open Access funding enabled and organized by CAUL and its Member Institutions This research was performed as part of the

Industrial Transformation Research Hub for Advanced Manufacturing of Medical Devices (IH150100024).

## Declarations

**Ethics approval** Not applicable.

**Consent for publication** The authors provide their consent to publish this article.

**Conflict of interest** The authors declare no conflict of interest.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Oh S, Cha J, Kim D, Jeong J (2020) Quality inspection of casting product using CAE and CNN. 4th International Conference on Imaging, Signal Processing and Communications (ICISPC). IEEE, pp 34–38
- Aguilar-Torres MA, Argüelles-Cruz AJ, Yáñez-Márquez C (2008) A real time artificial vision implementation for quality inspection of industrial products. IEEE Electronics, Robotics and Automotive Mechanics Conference (CERMA'08) 277–282
- Campbell SL, Gear CW (1995) The index of general nonlinear DAES. Numer Math 72(2):173–196
- Jia J (2009) A machine vision application for industrial assembly inspection. Second international conference on machine vision, IEEE 172–176
- Aisyah L (2012) Comparing the performance of ordinal logistic regression and artificial Neural network when analyzing ordinal data. Ph.D. thesis, Oklahoma State University
- Brosnan T, Sun D (2004) Improving quality inspection of food products by computer vision—a review. J Food Eng 61(1):3–16
- Utku H (2000) Application of the feature selection method to discriminate digitized wheat varieties. J Food Eng 46(3):211–216
- Sharma R, Kumar M, Alam M (2021) Image processing techniques to estimate weight and morphological parameters for selected wheat refractions. Sci Rep 11(1):1–12
- Panigrahi S, Misra MK, Willson S (1998) Evaluations of fractal geometry and invariant moments for shape classification of corn germplasm. Comput Electron Agric 20(1):1–20
- Li Y, Feng X, Liu Y, Han X (2021) Apple quality identification and classification by image processing based on convolutional neural networks. Sci Rep 11(1):1–15
- Xu J, Riccioli C, Sun D (2017) Comparison of hyperspectral imaging and computer vision for automatic differentiation of organically and conventionally farmed salmon. J Food Eng 196:170–182
- Sonego M, Fleck C, Pessan LA (2020) Hierarchical levels of organization of the brazil nut mesocarp. Sci Rep 10(1):1–13
- Sun D (2000) Inspecting pizza topping percentage and distribution by a computer vision method. J Food Eng 44(4):245–249

14. Tellaeche A., Robles B (2010) 3D machine vision and artificial neural networks for quality inspection in mass production pieces. IEEE 15th Conference on Emerging Technologies & Factory Automation 1–4
15. Xie H, Lu F, Ouyang G, Shang X, Zhao Z (2017) A rapid inspection method for encapsulating quality of pet bottles based on machine vision. 3rd IEEE International Conference on Computer and Communications (ICCC) 2025–2028
16. Bahaghighat M, Mirfattahi M, Akbari L, Babaie M (2018) Designing quality control system based on vision inspection in pharmaceutical product lines. IEEE International Conference on Computing, Mathematics and Engineering Technologies 1–4
17. García DF, Usamentiaga R et al (2021) Rail surface inspection system using differential topographic images. IEEE Trans Ind Appl 57(3):2994–3003
18. Li Y, Gao X, Zheng Q, Gao PP, Zhang N (2019) Weld cracks nondestructive testing based on magneto-optical imaging under alternating magnetic field excitation. Sensors and Actuators 285:289–299
19. Loh H-H, Lu M-S (1999) Printed circuit board inspection using image analysis. IEEE Trans Ind Appl 35(2):426–432
20. Yin X, Fang Z (2008) Vision based automatic needle inspection system for lancets. 3rd IEEE Conference on Industrial Electronics and Applications 1014–1018
21. Zhang Z et al (2021) Knitting needle fault detection system for hosiery machine based on laser detection and machine vision. Text Res J 91(1–2):143–151
22. Francois C (2018) Deep learning with Python. Manning Publications Company
23. Ko T et al (2017) Machine learning based anomaly detection via integration of manufacturing, inspection and after sales data. Ind Manag Data Syst
24. Laferrrière AM (2001) Method and system for determining and/or using illumination maps in rendering images. Google Patents, US Patent 6.226.005
25. Hata T et al (2012) Three- and four-dimensional HDlive rendering images of normal and abnormal fetuses. Arch Gynecol Obstet 286(6):1431–1435
26. McCullagh P (1980) Regression models for ordinal data. J Roy Stat Soc 42(2):109–127
27. Skapetis T, Doan-Tran PD, Hossain NM (2019) Evaluation of bevelled needle tip deformation with dental inferior alveolar nerve blocks. Aust Endod J 45(3):325–330
28. Gaitán-Fonseca C et al (2015) Bevel tip deformation in new and used dental needles. Microsc Res 3(01):1
29. Fisher RA (1936) Design of experiments. Br Med J 1(3923):554–554
30. Joe S, Kuo FY (2008) Constructing sobol sequences with better two-dimensional projections. SIAM J Sci Comput 30(5):2635–2654
31. Karaman S, Frazzoli E (2010) Incremental sampling-based algorithms for optimal motion planning. Robotics Science and Systems VI 104(2)
32. Anderson MJ, Whitcomb PJ (2000) Design of experiments. Kirk-Othmer Encycl Chem Technol 1–22
33. Aguilera A, Escabias M (2008) Solving multicollinearity in functional multinomial logit models for nominal and ordinal responses. Functional and Operatorial Stat 7–13
34. McCullagh P, Nelder JA (1989) Generalized linear models. Monogr Statist Appl Probab 37
35. Bender R, Grouven U (1998) Using binary logistic regression models for ordinal data with non-proportional odds. J Clin Epidemiol 51(10):809–816
36. Tranmer M, Elliot M (2008) Binary logistic regression. Cathie Marsh for Census and Survey Research 20
37. Hosmer S, Lemeshow D, Sturdivant R (2013) Applied logistic regression, vol 398. John Wiley & Sons
38. Minitab use'r guide 2: Data analysis and quality tools Release 13 (2000)
39. Widiharih T, Mukid M et al (2020) Probability of caused factors stroke disease use link and reliability functions. J Phys Conf Ser 1524(1):012124
40. Guan J, Zhao Y (2020) Parameter estimation approaches to tackling measurement error and multicollinearity in ordinal probit models. Communications in Statistics-Theory and Methods 49(16):3835–3859
41. Benesty J, Chen J, Huang Y, Cohen I (2009) Pearson correlation coefficient. Noise Reduction in Speech Processing 1–4
42. Niles HE (1922) Correlation, causation and wright's theory of path coefficients. Genetics 7(3):258–273
43. Li G, Doss H (1993) Generalized Pearson-Fisher chi-square goodness-of-fit tests, with applications to models with life history data. Ann Stat 772–797
44. Pulkstenis E, Robinson TJ (2004) Goodness-of-fit tests for ordinal response regression models. Stat Med 23(6):999–1014
45. Pan H, Cole T (2004) A comparison of goodness of fit tests for age related reference ranges. Stat Med 23:1749–1765
46. Hwang JP, Park S, Kim E (2011) A new weighted approach to imbalanced data classification problem via support vector machine with quadratic cost function. Expert Syst Appl 38(7):8580–8585
47. Keerthi SS, Lin C-J (2003) Asymptotic behaviors of support vector machines with gaussian kernel. Neural Comput 15(7):1667–1689
48. MathWorks. Classification models in automated classifier training. <https://au.mathworks.com/help/stats/train-classification-models-in-classification-learner-app.html>. Accessed 7 Oct 2021
49. Safavian SR, Landgrebe D (1991) A survey of decision tree classifier methodology. IEEE Trans Syst Man Cybern 21(3):660–674
50. Hollander M, Wolfe DA, Chicken E (2013) Nonparametric statistical methods, vol 751. John Wiley & Sons
51. Sachs L (2012) Applied statistics: a handbook of techniques. Springer Science
52. Haralick RM, Shapiro LG (1992) Computer and robot vision, vol 2. Addison-Wesley Reading

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.