



(51) International Patent Classification:
G06T 7/00 (2017.01)

(21) International Application Number:
PCT/US2024/033334

(22) International Filing Date:
11 June 2024 (11.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/507,491 12 June 2023 (12.06.2023) US
63/584,185 21 September 2023 (21.09.2023) US

(71) Applicant: **ELI LILLY AND COMPANY** [US/US]; Lilly
Corporate Center, Indianapolis, Indiana 46285 (US).

(72) Inventors: **CHUNG, Yu-Min**; c/o ELI LILLY AND COM-
PANY, P.O. Box 6288, Indianapolis, Indiana 46206-6288
(US). **COMPARINI, Edwin Rene**; c/o ELI LILLY
AND COMPANY, P.O. Box 6288, Indianapolis, Indi-
ana 46206-6288 (US). **ZHANG, Bo**; c/o ELI LILLY

AND COMPANY, P.O. Box 6288, Indianapolis, Indiana
46206-6288 (US).

(74) Agent: **SHUM, Arthur C.H.** et al.; ELI LILLY
AND COMPANY, P.O. Box 6288, Indianapolis, Indiana
46206-6288 (US).

(81) Designated States (*unless otherwise indicated, for every
kind of national protection available*): AE, AG, AL, AM,
AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ,
CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM,
DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT,
HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG,
KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,
MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA,
NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO,
RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH,
TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS,
ZA, ZM, ZW.

(84) Designated States (*unless otherwise indicated, for every
kind of regional protection available*): ARIPO (BW, CV,
GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST,
SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ,

(54) Title: METHODS AND SYSTEMS FOR DETECTING DEFECTS USING TOPOLOGICAL PERSISTENCE FEATURES

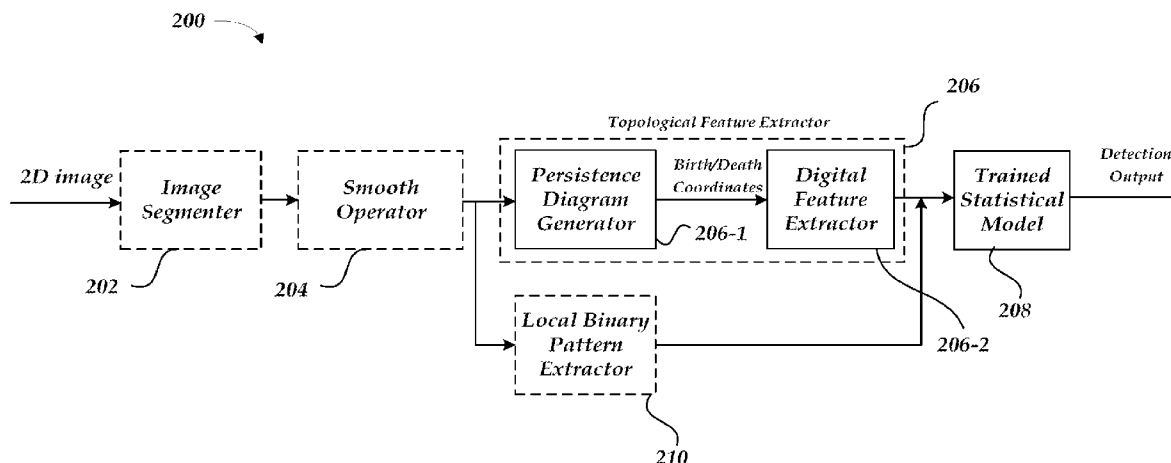


FIG. 2

(57) **Abstract:** The techniques described herein relate to computerized methods and apparatuses for detecting defects in manufactured components. In some embodiments, a system for monitoring for manufacturing defects may include an image capturing device configured to capture an image of a manufactured component. The system may compute a set of birth and death coordinates for each topological feature of a plurality of topological features in the image; determine a digital feature based on the computed sets of birth and death coordinates for each topological feature of the plurality of topological features; and determine, based on the digital feature, whether the image is indicative of a defect in the manufactured component. The sets of birth and death coordinates may be computed using topological data analysis. The digital feature may be used to determine whether the image is indicative of a defect in the manufactured component using a statistical model.

RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Declarations under Rule 4.17:

- *as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii))*
- *as to the applicant's entitlement to claim the priority of the earlier application (Rule 4.17(iii))*

Published:

- *with international search report (Art. 21(3))*
- *before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments (Rule 48.2(h))*
- *with information concerning incorporation by reference of missing element or part (Rule 20.6)*

METHODS AND SYSTEMS FOR DETECTING DEFECTS USING TOPOLOGICAL PERSISTENCE FEATURES

BACKGROUND

[0001] In a production assembly line, defect detection systems are used to automatically detect defects in manufactured components that move through the assembly line. Some defect detection systems capture an image of a manufactured component while the manufactured component enters into, is processed through, or exits the assembly line and process the captured image to determine whether the image indicates a defect in the manufactured component.

SUMMARY

[0002] The present disclosure relates to techniques for monitoring manufacturing defects in manufactured components. The techniques provide a computerized method for monitoring for manufacturing defects. The method may include: receiving a two-dimensional (2D) image of a manufactured component; computing a respective set of values indicating a life for each topological feature of a plurality of topological features in the 2D image; determining a digital feature based on the computed sets of values for the plurality of topological features; and determining, based on the digital feature, whether the 2D image is indicative of a defect in the manufactured component.

[0003] In some embodiments, each topological feature of the plurality of topological features for a 2D image may include a shape comprising an isolated block or a block in a first color completely enclosed by a region in a second color, wherein each pixel in the shape is below a filtration threshold. In some embodiments, the set of values for each topological feature of the plurality of topological features in the 2D image may include a set of birth and death coordinates each respectively representing a birth and a death of the topological feature associated with a filtration threshold. In some embodiments, in determining the digital feature, the method may determine a persistence statistics vector, a persistence curve vector, or a combination thereof, based on the computed sets of birth and death coordinates for the plurality of topological features. In some embodiments, in determining whether the 2D image is indicative of a defect in the manufactured component, the method may use a trained statistical model and the digital feature as an input to the trained statistical model. In some embodiments, the trained statistical model may be an XGBoost classifier.

[0004] In some embodiments, the method may further include segmenting the 2D image before computing the sets of values for the plurality of topological features. In some embodiments, the method may further include performing a smoothing operation, e.g., a morphological closing operation, on the 2D image before computing the sets of values for the plurality of topological features. In some embodiments, the method may further include determining a local binary pattern based on the 2D image, where using the trained statistical model and the digital feature comprises additionally using the local binary pattern as an input to the trained statistical model.

[0005] In some embodiments, receiving the 2D image of the manufactured component may include capturing the 2D image of the manufactured component while the manufactured component is moving along a conveyor belt. The manufactured component may be a syringe. Capturing the 2D image of the manufactured component may include capturing the 2D image from a top of the manufactured component. The defect may include a crack, a chip, or both.

BRIEF DESCRIPTION OF THE DRAWINGS

[0006] Additional embodiments of the disclosure, as well as features and advantages thereof, will become more apparent by reference to the description herein taken in conjunction with the accompanying drawings. The components in the figures are not necessarily to scale. Moreover, in the figures, like-referenced numerals designate corresponding parts throughout the different views.

[0007] FIG. 1 is a diagram of an exemplary system for detecting defects in manufactured components while the manufactured components are moving along a conveyor belt, according to some embodiments.

[0008] FIG. 2 is a diagram of an exemplary system for detecting whether a manufactured component in a 2D image has a defect, according to some embodiments.

[0009] FIG. 3 is a flow chart showing an exemplary computerized method for detecting manufactured component defects, according to some embodiments.

[0010] FIG. 4 is a flow chart showing an exemplary computerized method for determining a topological feature that is used for detecting whether a manufactured component has a defect, according to some embodiments.

[0011] FIG. 5 is a flow chart showing an exemplary computerized method for determining a local binary pattern vector that is used for detecting whether a manufactured component has a defect, according to some embodiments.

[0012] FIGS. 6A-6D show examples of 2D images containing a manufactured component that does not have any defect, according to some embodiments.

[0013] FIGS. 6E-6F show examples of 2D images containing a manufactured component that has a defect, according to some embodiments.

[0014] FIG. 7A shows an example of a 2D image containing a manufactured component that has a defect, according to some embodiments.

[0015] FIG. 7B shows an example local binary pattern of the 2D image as shown in FIG. 7A, according to some embodiments.

[0016] FIG. 7C shows an example histogram of the local binary pattern as shown in FIG. 7B, according to some embodiments.

[0017] FIG. 8 shows sets of birth/death coordinates for an example 2D image containing a manufactured component without and with morphological closing performed, according to some embodiments.

[0018] FIGS. 9A-9E show examples of topological features as the filtration threshold varies for a sample image, according to some embodiments.

[0019] FIG. 10A shows an example raw image and FIG. 10B shows an example of a segmented image from the raw image shown in FIG. 10A, according to some embodiments.

[0020] FIGS. 11A-11C show visualization of the spaces of features that represent the images absent defects and images having defects as detected by various configurations of the system, according to some embodiments.

[0021] FIG. 12 shows an illustrative implementation of a computer system that may be used to perform any of the aspects of the techniques and embodiments disclosed herein, according to some embodiments.

DETAILED DESCRIPTION

[0022] For the purposes of promoting an understanding of the principles of the present disclosure, reference will now be made to the embodiments illustrated in the drawings, and specific language will be used to describe the same. It will nevertheless be understood that no limitation of the scope of the invention is thereby intended.

[0023] In a pharmaceutical production line, for quality assurance purposes, a manufactured component, such as a syringe, a vial, or any other component, can be scanned for defect detection. This scanning may be done when the manufactured component enters an assembly

line, while the component is being processed through the assembly line, and/or when the component exits the assembly line.

[0024] The inventors have recognized and appreciated that conventional defect detection systems face challenges of achieving both speed and accuracy. For example, some conventional systems tend to be overly conservative with identifying defects and hence have high false positives, causing a waste/disposal of more samples than necessary. Some conventional systems, e.g., those using neural network models, may achieve a higher accuracy but at the cost of speed. For example, neural network based systems tend to require a large quantity of features and extensive computational resources, making them less suitable for real-time defect detection.

[0025] In a production assembly line, defect detection may need to be performed in real time so that the detection can be completed while a manufactured component is still on, or before exiting, the assembly line. In such manner, a defective manufactured component can be timely removed from the assembly line (e.g., before moving to the next manufacturing stage). In some embodiments, even if the defect is not identified until after the manufactured component moves to the next production stage or exits the assembly line, it would still be desirable to detect defects quickly so as to minimize the number of faulty components produced and/or processed. For example, for high speed assembly lines, defect detection may need to occur on the order of milliseconds. In an example assembly line, to maintain normal production rate, a defect detection system may only have around 150 milliseconds (ms) to respond, i.e. to determine whether a manufactured component (e.g., a syringe that has entered into the assembly line) has a defect or is damaged. Further, the determination as to whether there is a defect in the manufactured component typically needs to be sufficiently accurate in order to minimize false positives (which can cause unnecessary disposal of manufactured components with no defects) or false negatives (which can cause defective manufactured components to enter into the next product stage).

[0026] To solve the various technical problems discussed above or other issues, the inventors have recognized and appreciated that topological data analysis can be used for defect detection to achieve both high accuracy and fast execution speed. In particular, the inventors have appreciated that topological features may contain geometrical, shape, and/or texture characteristics associated with a manufactured component, which can be used to detect defects in manufactured components. The inventors have also recognized and appreciated that topological data analysis techniques may be suitable for extracting topological features

from an image of a manufactured component, where computing the topological features can be performed with high efficiency.

[0027] Accordingly, the inventors have developed new technologies for monitoring for manufacturing defects in manufactured components. Described herein are various techniques, including systems, computerized methods, and computer-executable instructions stored on non-transitory computer-readable media, that receive a two-dimensional (2D) image of a manufactured component. For example, the system may capture the 2D image of the manufactured component when the component is moving along a conveyor belt. In some examples, the 2D image of the manufactured component may be captured from the top of the manufactured component.

[0028] In some embodiments, the system may compute a set of values for each topological feature of a plurality of topological features in the 2D image, where the set of values indicate the life (e.g., start and end) of the topological feature. The plurality of topological features may contain geometrical, shape, and/or texture characteristics associated with a manufactured component. For example, the topological features may be location and rotation invariant. In some examples, a topological feature may be associated with a shape and may be of different levels. For example, a *0-th* level topological feature may include an isolated block in a first color (e.g., a connected group of at least one black pixel), and a *1-st* level topological feature may include a block in a second color enclosed by a region in the first color (e.g., a block of white pixels enclosed by black pixels that are connected to each other). For each of the topological features, the system may determine a set of values indicating the life of the topological feature, e.g., start and end of the topological feature associated with a varying threshold. In some embodiments, the set of values for a topological feature may include a set of birth and death coordinates, e.g., a pair of birth and death coordinates each indicating a value associated with a start and an end of the topological feature, respectively. In some embodiments, the birth and death coordinate pair for a topological feature are determined by varying a filtration threshold and comparing the intensities of the 2D image with the filtration threshold to identify the set of birth and death coordinates.

[0029] In some embodiments, the system may determine a digital feature based on the computed set of birth and death coordinates for each topological feature of the plurality of topological features. In some examples, the digital feature may include one or more persistence curve vectors based on the sets of birth and death coordinates for the plurality of topological features. For example, the digital feature may include a persistence statistics

vector, a persistence curve vector, or a combination thereof. In some embodiments, a persistence statistics vector may include one or more of a mean, a standard, a skewness, a kurtosis, or an entropy associated with the birth and death coordinates in the sets of birth/death coordinates. A persistence curve vector may also be determined based on the birth and death coordinates in the sets of birth and death coordinates.

[0030] In some embodiments, the system may determine, based on the digital feature, whether the 2D image is indicative of a defect in the manufactured component. In some embodiments, the system may use a trained statistical model and the digital feature as an input to the trained statistical model to detect the defect in the manufactured component. The statistical model may be any suitable model, such as a classifier, and may be implemented in any suitable framework, such as a gradient boosting machine learning model (XGBoost), a convolutional neural network (CNN), or any other suitable model. As a result, the output of the statistical model may indicate whether the manufactured component has a defect. Examples of defects that may be detected include a crack, a chip, and/or the like.

[0031] In some embodiments, additional features may also be provided to the statistical model to determine whether the 2D image is indicative of any defect in the manufactured component. For example, the system may determine a local binary pattern vector based on the 2D image and provide the local binary pattern vector as input to the statistical model. In some embodiments, the local binary pattern vector may be determined based on a local binary pattern image of the 2D image, where the local binary pattern image may contain grayscale and/or rotation invariant features. In some examples, the local binary pattern vector may be determined based on the histogram of the local binary pattern image.

[0032] In some embodiments, the captured 2D image of the manufactured component may be pre-processed to improve the performance of the topological feature extraction and subsequent detection. For example, the system may segment the 2D image (e.g., via cropping) before extracting the topological features. In some examples, the system may also perform smoothing operation over the 2D image, e.g., by performing a morphological closing operation. It is appreciated that other smoothing techniques may also be used.

[0033] The techniques described herein may provide advantages over conventional systems for detecting defects in manufactured components. For example, improved accuracy can be achieved over conventional systems with respect to measures, such as true positive, false negative, false positive, and true negative, area under receiver operating characteristic (AUROC), area under precision-recall curve (AUPRC), precision, recall, specificity and/or

the harmonic mean of precision and recall (F1). The techniques described herein can also achieve high speed that is sufficient to detect defects in real time. In one exemplary embodiment implemented on Applicant's manufacturing line for manufacturing and/or processing glass syringes, the processing speed for processing images collected under typical manufacturing operations show that 52ms can be achieved for processing each image, with slightly additional time for performing local binary pattern (33ms additionally) and morphological closing operation (0.6ms).

[0034] Whereas various embodiments have been described, it will be apparent to those of ordinary skill in the art that many more embodiments and implementations are possible. Accordingly, the embodiments described herein are examples, not the only possible embodiments and implementations. Furthermore, the advantages described above are not necessarily the only advantages, and it is not necessarily expected that all of the described advantages will be achieved with every embodiment.

[0035] FIG. 1 is a diagram of an exemplary system 100 for detecting defects in manufactured components while the manufactured components are moving along a conveyor belt, according to some embodiments. In some embodiments, system 100 may include an imaging capturing device (e.g., a camera 112) configured to obtain an image/video of a manufactured component 102 in a production line while the manufactured component 102 is moving along a conveyor belt 104. As shown, an example of the manufactured component 102 is a container. However, this is for exemplary purposes only, and it should be appreciated that the manufactured component can be any other device or component to be scanned for defect detection, such as a syringe, a glass vial, a bottle, a jar, and/or any suitable component. Although the manufactured component 102 is shown on a conveyor, the manufactured component may be moving on other platforms while being scanned/imaged, such as a turntable.

[0036] System 100 may further include a server 110 having at least a processor, a memory, a storage medium, and/or other components. The at least one processor may be configured to execute programming instructions stored in the memory to process the captured image/video. For example, the server 110 may analyze a 2D image in real time as each image is captured, and detect defect(s) in the image. In some examples, the inspection process for the manufactured component can be completed in a short period of time, such as a fraction of a second, during which time the system may determine whether the manufactured component may likely have a defect. Responsive to determining that the manufactured component may

likely have a defect, the system may perform one or more corrective measures, such as removing the manufactured component from the assembly line, before moving to the next production stage.

[0037] Alternatively, and/or additionally, responsive to determining that the manufactured component may likely have a defect, the system may generate an alert indicative of the detected potential defect. The alert may include any suitable form and may be communicated to a user device in any suitable communication protocols to indicate to the user that a defect is detected. For example, server 110 may send a notification, a warning, or a report to a user device 114, via a communication link (wired or wireless). The notification or the report may include information about the defect, and/or the type of defect. Upon receiving the alert at the user device 114, the operator may timely take corrective measures. Details of system 100 for detecting defects in manufactured components are further described with reference to FIGS. 2-10.

[0038] FIG. 2 is a diagram of an exemplary system 200 for detecting whether a manufactured component in a 2D image is likely to have a defect, according to some embodiments. In some embodiments, system 200 may be an example implementation of system 100 (FIG. 1). For example, system 200 may be implemented in server 110 (FIG. 1). System 200 may receive 2D images of manufactured components, where the 2D images may be captured in the manner as described in embodiments of FIG. 1. In non-limiting examples, the manufactured components may be syringes. Examples of defects that may be detected may include a crack, a chip, or both, or any other defects. FIGS. 6A-6F show example 2D images of a syringe that are captured from the top of the syringe. As shown, FIGS. 6A-6D are examples of 2D images containing a syringe that does not have any defect. FIGS. 6E-6F are examples of 2D images containing a syringe that has a defect.

[0039] Returning to FIG. 2, system 200 may include a topological feature extractor 206 configured to determine topological features from the 2D image. Topological feature extractor 206 may include a persistence diagram generator 206-1 and a digital feature extractor 206-2. The persistence diagram generator 206-1 may be configured to generate a persistence diagram of the 2D image. In some embodiments, the persistence diagram may be generated using topological data analysis (TDA) as will be explained in detail in FIGS. 3-4. The persistence diagram may include a plurality of sets of values, e.g., birth and death coordinates, where each set indicates a life of a topological feature of a plurality of topological features. For example, the set of values for a topological feature may include a

birth and death coordinate pair corresponding to the topological feature of a plurality of topological features. The birth and death coordinates in each pair may include two values (e.g., numerical values) indicative of the life (e.g., birth, death) of each topological feature of a plurality of topological features. For example, the birth and death coordinates of a topological feature may respectively indicate the birth and death of the topological feature. The plurality of topological features may contain geometrical, shape, and/or texture characteristics associated with a manufactured component. For example, the topological features may be location and rotation invariant.

[0040] With further reference to FIG. 2, the digital feature extractor 206-2 may be configured to convert the persistence diagram (e.g., birth and death coordinates) to a digital feature for use with subsequent statistical model 208. The digital feature output from the digital feature extractor 206-2 may include one or more persistence curve vectors that are determined based on the sets of birth/death coordinates for the plurality of topological features. For example, the digital feature may include a persistence statistics vector, a persistence curve vector, or a combination thereof. A persistence statistics vector may include one or more of mean, standard, skewness, kurtosis, or entropy associated with the birth and death coordinates pairs in the sets of birth/death coordinates. A persistence curve vector may also be determined based on the birth and death coordinates pairs in the sets of birth/death coordinates.

[0041] In some embodiments, system 200 may further include a trained statistical model 208 configured to determine, based on the digital feature from the topological feature extractor 206, whether the 2D image is indicative of a defect in the manufactured component. The statistical model may be any suitable model, such as a classifier, and may be implemented in any suitable framework, such as a gradient boosting machine learning model (XGBoost), a convolutional neural network (CNN), or any other suitable model. As a result, the output of the statistical model may indicate whether the manufactured component has any defect.

[0042] With further reference to FIG. 2, alternatively and/or additionally, system 200 may include a local binary pattern (LBP) extractor 210 configured to provide additional features to the trained statistical model 208. For example, the LBP extractor 210 may determine a local binary pattern vector based on the 2D image and provide the local binary pattern vector as input to the statistical model 208. In some embodiments, the local binary pattern vector may be determined based on local binary pattern image of the 2D image, where the local binary pattern image may be indicative of grayscale and/or rotation invariant features. In some

examples, the local binary pattern vector may be determined based on the histogram of the local binary pattern image.

[0043] With further reference to FIG. 2, system 200 may optionally include one or more pre-processing units to pre-process the 2D images. In some embodiments, system 200 may include an image segmenter 202 configured to segment (e.g., via cropping) the 2D image before extracting the topological features from the 2D image. Additionally and/or alternatively, system 200 may include a smooth operator 204 configured to further perform smoothing operation over the 2D image. For example, the smooth operator 204 may be a morphological closing operator. In some examples, the smooth operator 204 may be a Gaussian filter, a median filter, a bilateral filter, a morphological opening operator, or a combination thereof. It is appreciated that other suitable smoothing techniques may also be used. Having described systems 100 (FIG. 1) and 200 (FIG. 2), details of methods that may be implemented in systems 100 and 200 for detecting defects in manufactured components are further described with reference to FIGS. 3-10.

[0044] FIG. 3 is a flow chart showing an exemplary computerized method 300 for detecting manufactured component defects, according to some embodiments. In some embodiments, method 300 may be implemented in system 100 (e.g., server 110, FIG. 1) or system 200 (FIG. 2). For example, each act in method 300 may be implemented in a respective block in system 200. In non-limiting examples, method 300 may start with receiving a 2D image of a manufactured component, at act 302, as described herein above. For example, the 2D images may be captured by a camera, which is configured to capture a 2D image of a manufactured component while the manufactured component is moving along a moving platform, e.g., a conveyor.

[0045] In some embodiments, method 300 may further include determining topological features, at act 308. In some embodiments, act 308 may be performed by the topological feature extractor 206 in FIG. 2. Act 308 is further described in detail with reference to an exemplary process in FIG. 4. FIG. 4 is a flow chart showing an exemplary computerized method 400 for determining a topological feature that is used for detecting whether a manufactured component is likely to have a defect, according to some embodiments. In some embodiments, method 400 may be implemented in system 100, 200, such as topological feature extractor 206 (FIG. 2). As shown in FIG. 4, method 400 starts with computing birth and death coordinates for each topological feature in the 2D image at act 402, where the birth and death coordinates may include a plurality of birth and death coordinates pairs, expressed

as (b, d) , for each topological feature. In some embodiments, act 402 may be implemented in persistence diagram generator 206-1 (FIG. 2). Thus, the birth and death coordinates also make up the persistence diagram.

[0046] In some examples, a persistence diagram may include one or more sublevel persistence diagrams each corresponding to a topological feature. For example, for a 2D image, there may be two sublevel persistence diagrams: a 0 -th level persistence diagram (D_0) and 1 -st level persistence diagram (D_1). In some embodiments, the persistence diagram D_0 may contain a first set of birth death coordinates pairs $(b, d) \in D_0$ associated with a first type of topological feature (e.g., 0 -th level topological features), such as an isolated region in a first color (e.g., a black region having black pixels connected to each other). The persistence diagram D_1 may contain a second set of birth death coordinates pairs $(b, d) \in D_1$ associated a second type of topological feature (e.g., 1 -st level topological features), such as a block in a second color enclosed by a region in the first color (e.g., a block of white pixels completely enclosed by black pixels that are connected to each other). The first color and the second color (e.g., black and white) may be associated with a binarization operation for which a filtration threshold may be applied. Detailed operations for determining persistence diagrams D_0 and D_1 are further described with an example shown in FIGS. 9A-9E.

[0047] FIGS. 9A-9E show examples of topological features as a filtration threshold varies for a sample image, according to some embodiments. FIG. 9A illustrates a sample grayscale image having 3×3 pixels with the intensity value of each pixel shown. FIGS. 9B-9E show binarized images when a respective filtration threshold is applied to the original image shown in FIG. 9A. A filtration threshold may be varied between the minimum and maximum intensity values of the image $y(x)$, e.g., 1-10 in this example. When the filtration threshold increases by value from 1-10, the system tracks the changes in homology. Applying the filtration threshold to the image may result in a binarized image having two values: black and white, where black represents a pixel with a value at or below the threshold and white represents a pixel with a value above the threshold.

[0048] As shown in FIG. 9B, when a filtration threshold of value 1 is applied, a topological feature 902 appears (is born) from the binarized image. Topological feature 902 having an isolated region in black may thus be associated with a 0 -th level persistence diagram. As shown in FIG. 9C, when a filtration threshold of value 2 is applied, topological feature 904 is born from the binarized image. As shown, topological feature 904 is of the same type as the

topological feature 902 and may also be associated with the 0 -th level persistence diagram (e.g., D_0).

[0049] As shown in FIG. 9D, when a filtration threshold of value 3 is applied, topological features 902, 904 are merged. In some embodiments, when two topological features overlap, they are merged using an elder rule, under which the senior topological feature that was born first survives whereas the younger one disappears (dies). In the example in FIG. 9D, topological feature 902 survives and topological feature 904 dies. The merged region of two topological features comprises pixels from both topological features being merged. For example, in FIG. 9D, topological features 902, 904 are merged as topological feature 902 having the black region with a white opening therein.

[0050] With further reference to FIG. 9D, a new type of topological feature 906, which corresponds to a white block completely enclosed by a black region (e.g., a white region completely enclosed by black pixels), is also born. Thus, the topological feature 906 may be associated with a 1 -st level persistence diagram (e.g., D_1). As shown in FIG. 9E, when a filtration threshold of value 10 is applied, the entire image becomes black, causing topological feature 906 to die, with topological feature 902 surviving.

[0051] In some embodiments, the birth coordinate for a particular topological feature is the filtration threshold level t at which that particular topological feature comes into existence as the filtration threshold level is increased from the minimum value of $y(x)$ to the maximum value of $y(x)$. Similarly, the death coordinate for a particular topological feature is the filtration threshold level t at which that particular topological feature expires. A particular topological feature is considered to expire when it merges with another topological feature having an earlier birth coordinate, as discussed above. In the example in FIGS. 9A-9E, the extracted persistence diagrams include: $D_0 = \{(1, \infty), (2, 3)\}$ having two birth and death coordinates pairs and $D_1 = \{(3, 10)\}$ having one birth and death coordinates pair. In this example, the birth and death coordinates pairs $(1, \infty)$, $(2, 3)$ in D_0 respectively correspond to the topological features 902 and 904, whereas the birth and death coordinates pair $(3, 10)$ corresponds to the topological feature 906.

[0052] Alternatively, and/or additionally, persistence diagrams may be determined based on Betti numbers associated with a plurality of filtration threshold levels in a grayscale image. In some embodiments, a Betti number may be determined for each resulting binary image when applying a respective filtration threshold to the grayscale image. For example, at each filtration threshold value, a Betti number may represent the count of the number of

topological spaces (shapes) based on the pixel connectivity in the binary image. For a two-dimensional grayscale image, a Betti number may include a 0 -th level number and a 1 -st level number, respectively, representing the number of isolated regions (e.g., connected black regions) and the number of voids (e.g., white regions that are completely enclosed by black pixels) at a given filtration threshold value. In the examples in FIGS. 9B-9E, the Betti numbers are respectively (1, 0), (2,0), (1,1), and (1, 0), where the first number in each pair represents the 0 -th level Betti number, and the second number in each pair represents the 1 -st level Betti number.

[0053] In some embodiments, the birth and death coordinates pairs for persistence diagrams D_0, D_1 may be associated with the change of Betti numbers as the filtration threshold changes. For example, from FIG. 9B to 9C, the Betti number changes from (1, 0) to (2, 0), representing a birth of a new topological feature (e.g., the 0 -th Betti number changes from 1 to 2). From FIG. 9C to 9D, the Betti number changes from (2, 0) to (1, 1), representing a merge of topological features (e.g., the 0 -th Betti number changes from 2 to 1) and a birth of a new topological feature (e.g., the 1 -st Betti number changes from 0 to 1). As such, the topological features may also be associated with Betti numbers or changes of Betti numbers associated with filtration threshold levels.

[0054] Returning to FIG. 4, method 400 may further determine a digital feature (e.g., at acts 404-410) based on the birth and death coordinates. In some embodiments, acts 404-410 may be implemented in digital feature extractor 206-2 (FIG. 2). The digital feature is further described.

[0055] In some embodiments, method 400 may first determine midlife persistence values, at act 404. Method 400 may also determine lifespan persistence values, at act 406. The midlife persistence and lifespan persistence values may be determined for each set of birth and death coordinates obtained from act 402. In some embodiments, act 404 may calculate a lifespan persistence (L) for each respective topological feature of the plurality of topological features by subtracting the birth coordinate for that respective topological feature from the death coordinate for that respective topological feature. Act 406 may calculate a midlife persistence (M) for each respective topological feature in the plurality of topological feature by calculating a mean average of the set of birth and death coordinates for that respective topological feature.

[0056] With further reference to FIG. 4, method 400 may include determining a persistence statistics vector at act 408. In some embodiments, act 408 may calculate a first persistence

statistics vector based on one or more of a mean, a standard deviation, a skewness, a kurtosis, or an entropy of the calculated lifespan persistence (L) across all topological features in the 2D image. Alternatively, and/or additionally, act 408 may calculate a second persistence statistics vector based on one or more of a mean, a standard deviation, a skewness, a kurtosis, or an entropy of the calculated midlife persistence (M) across all topological features in the 2D image. Let $D = \{(b, d)\}$ be a persistence diagram. Consider $M := \{d + b, (b, d) \in D\}$ and $L := \{d - b, (b, d) \in D\}$. Persistence statistics may include summary statistics of M and L which are mean, standard deviation, skewness, kurtosis, and entropy for M and L .

[0057] With further reference to FIG. 4, method 400 may include determining a persistence curve vector, at act 410. In some embodiments, a persistence curve vector may be calculated based on the lifespan values of the sets of birth and death coordinates for the plurality of topological features. For example, a persistence curve vector may be expressed as a product of two persistence curves, each based on a sum of lifespan values of the sets of birth and death coordinates corresponding to each of the persistence diagrams (e.g., D_0, D_1). In a non-limiting example:

$$P(D)(x) := - \sum_{(b,d) \in D_x} \frac{d-b}{L_1(D)} \log \left(\sum_{(b,d) \in D_x} \frac{d-b}{L_1(D)} \right),$$

where $L_1(D) = \sum_{(b,d) \in D} (d-b)$ and $D_x = \{(b, d) \in D \mid b \leq x, d > x\}$. As shown, the second persistence curve in the persistence curve vector is the logarithm of the first persistence curve.

[0058] In some embodiments, the persistence curve vectors described above can have any suitable number of features (elements). In a non-limiting example, for a given persistence diagram D (e.g., D_0 or D_1 in the above example), the persistence statistics vector may include a 1D vector having 10 features (e.g., one feature each for the mean, standard deviation, skewness, kurtosis, and entropy for both midlife persistence values and lifespan persistence values of the topological features in persistence diagram D). In a non-limiting example, for a given D , the persistence curve vector may be a 256-feature vector (this example assumes x , which represents intensity in the received 2D image of the manufactured component, may vary between 256 different levels from lowest to highest intensity. If the 2D image may vary between a greater number of intensity levels, e.g., between 512 or 1024 levels, the persistence curve vector may be a 512 or 1024-feature vector, respectively). Thus, for two persistence diagrams D_0 and D_1 , the digital feature may be a vector having $(10 + 256) \times 2 = 532$

elements. It is appreciated that the number of features in the digital feature can vary. It is further appreciated that the number of features in the digital features may vary by the number of persistence diagrams (the number of topological features) being extracted from the 2D image. For example, whereas the above example shows two persistence diagrams corresponding to two topological features (e.g., D_0 or D_1), more or fewer than two levels of persistence diagrams may be extracted and, as a result, the number of features in the digital features may also vary.

[0059] Returning to FIG. 3, method 300 may further include using the topological feature to determine whether the 2D image is indicative of a defect in the manufactured component, at act 312. In some embodiments, the topological feature used for act 312 may be the digital feature obtained from act 308 (details shown in FIG. 4). In some embodiments, in determining whether the 2D image is indicative of a defect in the manufactured component, act 312 may use a trained statistical model. Any suitable statistical model may be used. For example, act 312 may use a gradient boosting machine learning model (XGBoost), a convolutional neural network (CNN), or any other suitable model. Such a model may be trained using digital features similar to those obtained from act 308 in combination with ground truth labelling data that indicates whether a particular digital feature is associated with an image of a defective or not-defective component.

[0060] With further reference to FIG. 3, method 300 may optionally include one or more additional acts. For example, method 300 may optionally include segmenting the 2D image, at act 304, before determining the topological feature (act 308). In some embodiments, act 304 may be implemented in the image segmenter 202 (FIG. 2). In some embodiments, segmenting the 2D image may include cropping. In some embodiments, cropping may center the manufactured component in the image and/or trim off some background to reduce the computation required to process the image. In some embodiment, method 300 may impose a uniform boundary across all images. For example, let $R = (r_{i,j}) \in R^{640 \times 480}$ be the matrix of raw image. When cropping, pixels (i, j) that satisfy the range of $100 \leq i \leq 400$ and $200 \leq j \leq 500$ are kept while the pixels beyond that range are removed. In some embodiments, the cropping may only need to determine a rough boundary to include the manufactured component because the persistence diagrams are translation and rotation invariant. In other words, the cropping does not need to obtain the perfect location for the manufactured component. As an example, FIG. 10B shows an example of a segmented image from the raw image shown in FIG. 10A.

[0061] With further reference to FIG. 3, method 300 may optionally include performing a smoothing operation at act 306. In some embodiments, act 306 may be implemented in the smooth operator 204 (FIG. 2). Smoothing operation may suppress noise in the 2D image. Various known image processing technologies may be used to perform smoothing operation over the image. In a non-limiting example, the smooth operation may include a morphological closing operation. Any suitable size structuring element (e.g., 3x3) may be used in the morphological closing operation. It is appreciated that other smoothing techniques may also be used.

[0062] The inventors have appreciated and recognized that an image after smoothing may exhibit fewer topological features (e.g., those caused by noise are removed). This results in performance improvement of the statistical model (e.g., 208 in FIG. 2). In a non-limiting example, FIG. 8 shows the sets of birth/death coordinates for an example 2D image containing a manufactured component without and with morphological closing performed. FIG. 8(a) shows the original 2D image, wherein FIG. 8(d) shows the smoothed image of the image in FIG. 8(a) after performing the morphological closing operation. FIGS. 8(b) and 8(c) respectively show the 0 -th level persistence diagram (D_0) and 1 -st level persistence diagram (D_1) each containing sets of birth (horizontal axis) and death (vertical axis) coordinates. Similarly, FIGS. 8(e) and 8(f) respectively show the 0 -th level and 1 -st level persistence diagrams for the smoothed image in FIG. 8(d). As shown, relative fewer birth and death coordinates are extracted from the smoothed image.

[0063] Returning to FIG. 3, method 300 may optionally determine a local binary pattern, at act 310, and provide the local binary pattern as additional features to the statistical model (e.g., act 312, or 208 in FIG. 2). In some embodiments, act 310 may be implemented in local binary pattern extractor 210 (FIG. 2). The details of extracting the local binary pattern are further described with reference to FIG. 5, which is a flow chart showing an exemplary computerized method 500 for determining a local binary pattern vector that is used for detecting whether a manufactured component has defect, according to some embodiments.

[0064] With reference to FIG. 5, in some embodiments, method 500 may start with determining local binary pattern from the 2D image, at act 502. In some examples, act 502 may be configured to extract uniform local binary pattern, which may be grayscale and rotation invariant. Method 500 may also determine a histogram of the local binary pattern, at act 504 and determine a local binary pattern vector based on the histogram, at act 506. The operations of various acts in method 500 are further illustrated. FIG. 7A shows an example of

2D image containing a manufactured component that has defect, according to some embodiments. FIG. 7B shows an example local binary pattern of 2D image as shown in FIG. 7A, according to some embodiments. In this example, the local binary pattern from act 502 may be a grayscale image. In the example shown, the grayscale in the local binary pattern may range from 0-24. It is appreciated that any other suitable range may be possible.

[0065] FIG. 7C shows an example histogram of a local binary pattern as shown in FIG. 7B, according to some embodiments. In the example shown, the size of the histogram corresponds to the grayscale range in the local binary pattern, e.g., 25 (0-24). As such, act 506 may determine a local binary pattern vector based on the histogram, where the size of the vector is the size of the histogram, and the values of each element in the vector corresponds to the histogram value at each bin. In the example shown, the local binary pattern vector has a size of 25. It is appreciated that any other suitable size is possible for the local binary pattern vector.

[0066] Returning to FIGS. 2-3, having described various blocks and methods that can be implemented in system 200, the digital feature provided to the trained statistical model may include a 1D vector of any suitable size. As described in various examples herein, the digital feature may concatenate the persistence curve vector from topological feature extractor 206 and/or the local binary pattern vector from local binary pattern extractor 210. In a non-limiting example, when using both the topological feature extractor 206 and the local binary pattern extractor 210, according to the examples described above, the size of the persistence curve vector is 532, the size of the local binary pattern vector is 25, resulting in the total size of the digital feature being 557. As a result, the 2D image is compressed in an efficient manner and converted to a 1D vector of 557 elements.

[0067] Various embodiments described in the present disclosure have been implemented and tested in various configurations. For example, with reference to FIG. 2, in a first configuration, image segmenter 202, topological feature extractor 206, and trained statistical model 208 were implemented without using smooth operator 204 or local binary pattern extractor 210. In a second configuration, image segmenter 202, smooth operator 204, topological feature extractor 206, and trained statistical model 208 were implemented without using local binary pattern extractor 210. In a third configuration, all of the components (e.g., image segmenter 202, smooth operator 204, topological feature extractor 206, local binary pattern extractor 210, and trained statistical model 208) were implemented. The experimental results for these configurations are further described below.

[0068] Table 1 shows the validation result of the system for the different configurations as described above, using a validation data set. Validation can be used to fine-tune the model parameters. TN, FP, FN, and TP represent true negatives, false positives, false negatives, and true positives, respectively. $F1 = 2 / (\text{Recall}^{-1} + \text{Precision}^{-1})$, where $\text{Recall} = TP / (TP + FN)$, and $\text{Precision} = TP / (TP + FP)$.

Configurations	TN	FP	FN	TP	AUROC	F1
First	10741	20	43	1123	0.981	0.973
Second	10748	13	40	1126	0.982	0.977
Third _	10752	9	31	1135	0.986	0.983

[0069] Table 2 shows the testing result of the system for different configurations as described above, using a testing data set. Testing can be used to evaluate the performance of the model once the parameters are fine-tuned.

Configurations	TN	FP	FN	TP	AUROC	F1
First	7539	24	28	828	0.982	0.970
Second	7546	17	31	825	0.981	0.972
Third _	7554	9	26	830	0.984	0.979

In both validation and testing of the system, parameters in XGBoost, such as learning rate, n estimators, and scale pos weight were respectively 0.1, 100, and 1. The split for training, validation, and testing data was 65-20-15. As shown in Tables 1 and 2, the three configurations all yielded classification performance with AUROC values exceeding 0.98. Further, with the complexity of the system increasing from the first configuration to the third configuration, the performance of the system improved, e.g., both FP and FN decreased as shown in Tables 1 and 2.

[0070] Table 3 shows the results on holdout data sets (which are obtained from production line on different days and not used in the training).

Configurations	Holdout	TN	FP	FN	TP	Holdout	TN	FP	FN	TP
First	0304	10249	61	4	33	0305	11291	44	5	43
Second	0304	10253	57	2	35	0305	11319	16	9	39
Third _	0304	10284	26	1	36	0305	11313	22	7	41

The holdout dataset generated similar results to those for validation and testing. Further, similar to the results shown in Tables 1 and 2, the false positives (FP) and false negatives (FN) for the holdout data sets also decreased with the configurations changing from the first to the third configurations with increased complexity. Such results suggest that the techniques provided in the present disclosure perform adequately on unseen datasets.

[0071] FIGS. 11A-11C show visualization of the spaces of features that represent images absent defects (good images) and images having defects (defect images) as detected by various configurations of the system, in 3D space. The features such as described above with respect to the persistence curve vector, and/or local binary pattern vector, are projected into the 3D space using a suitable projection method. For example, FIGS. 11A and 11B show visualizations of the 3D spaces onto which 532 features (e.g., persistence curve vector) are projected, where defect images were respectively detected by the first and second configurations of the system described with respect to Tables 1-3. Similarly, FIG. 11C shows visualization of the 3D space onto which 557 features (e.g., persistence curve vector and local binary pattern vector) are projected, where defect images were detected by the third configuration of the system. In these drawings, the dots in different shades/color represent images absent defects and having defects as detected by the system. As shown in FIGS. 11A-11C, separations between these two classes of images can be seen.

[0072] In testing the various embodiments, it is shown that the techniques as described (e.g., various configurations in FIG. 2) can be performed reasonably efficiently in a modern computer, such as a CPU architecture with Intel x86-64 core having multiple subsystems. In some embodiments, the various embodiments or components of the system (e.g., those shown and discussed in conjunction with FIG. 2) may be configured such that the software can be optimized for execution on multiple instruction execution units (multi-threading), a multi-level cache hierarchy, multi-stage instruction pipelines, and/or branch prediction units. For example, XGBoost may be implemented in a statistical model (e.g., statistical model 208 in FIG. 2) that implements a tree boosting algorithm. Such an algorithm may be configured to build multiple parallel, data-independent trees for computing partial predictions, which can be combined to generate the final prediction results. Each tree may have a simple binary tree structure with a limited number of levels (e.g., 10 levels), and can be executed in parallel in a multi-threaded CPU. In some embodiments, the XGBoost algorithm may be configured to use exact greedy algorithm to find the optimal possible splits of all the features. The splits may indicate where the tree traversing procedure branches to the next tree leaf, where such tree branching strategies may be used to utilize the CPU's branch prediction unit.

[0073] In some embodiments, the various components of the system disclosed herein may also be optimized to utilize the memory efficiently. For example, XGBoost may be configured to map the feature data (e.g., topological features) to memory structures in order to increase the CPU cache utilization. For example, the feature data may be stored in in-

memory units (e.g., blocks). In some embodiments, data may be represented in Compressed Column format. Blocks are then sorted, which turns a data finding procedure into a linear scan. Since CPU cache stores a complete cache line, a linear scan in such a scenario will yield an optimal cache hit rate. In some embodiments, for non-continuous memory accesses, XGBoost may be configured to leverage a cache-aware prefetching algorithm. For example, an internal buffer may be allocated by each thread, to where XGBoost prefetches and performs accumulation over gradients.

[0074] FIG. 12 shows an illustrative implementation of a computer system that may be used to perform any of the aspects of the techniques and embodiments disclosed herein, according to some embodiments. An illustrative implementation of a computer system 1000 that may be used to perform any of the aspects of the techniques and embodiments disclosed herein is shown in FIG. 12. For example, the computer system 1000 may be installed in system 100 of FIG. 1, such as by server 110. The computer system 1000 may be configured to perform various methods and acts as described in FIGS. 3-5. The computer system 1000 may include one or more processors 1010 and one or more non-transitory computer-readable storage media (e.g., memory 1020 and one or more non-volatile storage media 1030) and a display 1040. The processor 1010 may control writing data to and reading data from the memory 1020 and the non-volatile storage device 1030 in any suitable manner, as the aspects of the invention described herein are not limited in this respect. In some embodiments, the computer system 1000 may also be a complete system on module (SOM), such as NVIDIA's Jetson module, which includes CPU, GPU, memory, and other components in a system. In some embodiments, the computer system 1000 may be located at any suitable site. For example, the computer system 1000 (e.g., server 110) may be co-located with the product line having the conveyor belt 104 or may be on the network. In other variations, the system may not need to include a memory, but instead programming instructions are running on one or more virtual machines or one or more containers on a cloud. For example, the various methods illustrated above may be implemented by a server on a cloud that includes multiple virtual machines, each virtual machine having an operating system, a virtual disk, virtual network and applications, and the programming instructions for detecting defects in manufactured components may be stored on one or more of those virtual machines on the cloud.

[0075] To perform functionality and/or techniques described herein, the processor 1010 may execute one or more instructions stored in one or more computer-readable storage media

(e.g., the memory 1020, storage media, etc.), which may serve as non-transitory computer-readable storage media storing instructions for execution by the processor 1010.

[0076] In connection with techniques described herein, code used to, for example, detect defects in manufactured components may be stored on one or more computer-readable storage media of computer system 1000. Processor 1010 may execute any such code to provide any techniques for detecting defects as described herein. Any other software, programs or instructions described herein may also be stored and executed by computer system 1000. It will be appreciated that computer code may be applied to any aspects of methods and techniques described herein. For example, computer code may be applied to interact with an operating system to detect defects through conventional operating system processes.

[0077] The various methods or processes outlined herein may be coded as software that is executable on one or more processors that employ any one of a variety of operating systems or platforms. Additionally, such software may be written using any of numerous suitable programming languages and/or programming or scripting tools, and also may be compiled as executable machine language code or intermediate code that is executed on a virtual machine or a suitable framework.

[0078] In this respect, various inventive concepts may be embodied as at least one non-transitory computer readable storage medium (e.g., a computer memory, one or more floppy discs, compact discs, optical discs, magnetic tapes, flash memories, circuit configurations in Field Programmable Gate Arrays or other semiconductor devices, etc.) encoded with one or more programs that, when executed on one or more computers or other processors, implement the various embodiments of the present invention. The non-transitory computer-readable medium or media may be transportable, such that the program or programs stored thereon may be loaded onto any computer resource to implement various aspects of the present invention as discussed above.

[0079] The terms “program,” “software,” and/or “application” are used herein in a generic sense to refer to any type of computer code or set of computer-executable instructions that can be employed to program a computer or other processor to implement various aspects of embodiments as discussed above. Additionally, it should be appreciated that according to one aspect, one or more computer programs that when executed perform methods of the present invention need not reside on a single computer or processor, but may be distributed in

a modular fashion among different computers or processors to implement various aspects of the present invention.

[0080] Computer-executable instructions may be in many forms, such as program modules, executed by one or more computers or other devices. Generally, program modules include routines, programs, objects, components, data structures, etc. that perform particular tasks or implement particular abstract data types. Typically, the functionality of the program modules may be combined or distributed as desired in various embodiments.

[0081] Also, data structures may be stored in non-transitory computer-readable storage media in any suitable form. Data structures may have fields that are related through location in the data structure. Such relationships may likewise be achieved by assigning storage for the fields with locations in a non-transitory computer-readable medium that convey relationship between the fields. However, any suitable mechanism may be used to establish relationships among information in fields of a data structure, including through the use of pointers, tags or other mechanisms that establish relationships among data elements.

[0082] Various inventive concepts may be embodied as one or more methods, of which examples have been provided. The acts performed as part of a method may be ordered in any suitable way. Accordingly, embodiments may be constructed in which acts are performed in an order different than illustrated, which may include performing some acts simultaneously, even though shown as sequential acts in illustrative embodiments.

[0083] The indefinite articles “a” and “an,” as used herein in the specification and in the claims, unless clearly indicated to the contrary, should be understood to mean “at least one.” As used herein in the specification and in the claims, the phrase “at least one,” in reference to a list of one or more elements, should be understood to mean at least one element selected from any one or more of the elements in the list of elements, but not necessarily including at least one of each and every element specifically listed within the list of elements and not excluding any combinations of elements in the list of elements. This allows elements to optionally be present other than the elements specifically identified within the list of elements to which the phrase “at least one” refers, whether related or unrelated to those elements specifically identified.

[0084] The phrase “and/or,” as used herein in the specification and in the claims, should be understood to mean “either or both” of the elements so conjoined, i.e., elements that are conjunctively present in some cases and disjunctively present in other cases. Multiple elements listed with “and/or” should be construed in the same fashion, i.e., “one or more” of

the elements so conjoined. Other elements may optionally be present other than the elements specifically identified by the “and/or” clause, whether related or unrelated to those elements specifically identified. Thus, as a non-limiting example, a reference to “A and/or B”, when used in conjunction with open-ended language such as “comprising” can refer, in one embodiment, to A only (optionally including elements other than B); in another embodiment, to B only (optionally including elements other than A); in yet another embodiment, to both A and B (optionally including other elements); etc.

[0085] As used herein in the specification and in the claims, “or” should be understood to have the same meaning as “and/or” as defined above. For example, when separating items in a list, “or” or “and/or” shall be interpreted as being inclusive, i.e., the inclusion of at least one, but also including more than one, of a number or list of elements, and, optionally, additional unlisted items. Only terms clearly indicated to the contrary, such as “only one of” or “exactly one of,” or, when used in the claims, “consisting of,” will refer to the inclusion of exactly one element of a number or list of elements. In general, the term “or” as used herein shall only be interpreted as indicating exclusive alternatives (i.e. “one or the other but not both”) when preceded by terms of exclusivity, such as “either,” “one of,” “only one of,” or “exactly one of.” “Consisting essentially of,” when used in the claims, shall have its ordinary meaning as used in the field of patent law.

[0086] Use of ordinal terms such as “first,” “second,” “third,” etc., in the claims to modify a claim element does not by itself connote any priority, precedence, or order of one claim element over another or the temporal order in which acts of a method are performed. Such terms are used merely as labels to distinguish one claim element having a certain name from another element having a same name (but for use of the ordinal term).

[0087] The phraseology and terminology used herein is for the purpose of description and should not be regarded as limiting. The use of “including,” “comprising,” “having,” “containing,” “involving,” and variations thereof, is meant to encompass the items listed thereafter and additional items.

[0088] Having described several embodiments of the invention in detail, various modifications and improvements will readily occur to those skilled in the art. Such modifications and improvements are intended to be within the spirit and scope of the invention. Accordingly, the foregoing description is by way of example only, and is not intended as limiting.

[0089] Various aspects are described in this disclosure, which include, but are not limited to, the following aspects:

[0090] (1) A computerized method for monitoring for manufacturing defects, the method comprising: receiving a two-dimensional (2D) image of a manufactured component; computing a respective set of values indicating a life for each topological feature of a plurality of topological features in the 2D image; determining a digital feature based on the computed sets of values for the plurality of topological features; and determining, based on the digital feature, whether the 2D image is indicative of a defect in the manufactured component.

[0091] (2) The method of aspect 1, wherein each topological feature of the plurality of topological features comprises a shape comprising an isolated block or a block in a first color completely enclosed by a region in a second color, wherein each pixel in the shape is below a filtration threshold.

[0092] (3) The method of aspect 1, wherein the set of values for each topological feature of the plurality of topological features in the 2D image include a set of birth and death coordinates each respectively representing a birth and a death of the topological feature associated with a filtration threshold. Additionally, and/or alternatively, the topological features may correspond to Betti numbers or change of Betti numbers associated with filtration threshold levels.

[0093] (4) The method of aspect 3, wherein determining the digital feature comprises: determining a persistence statistics vector, a persistence curve vector, or a combination thereof, based on the computed sets of birth and death coordinates for the plurality of topological features.

[0094] (5) The method of aspect 4, further comprising: determining a first set of values, wherein each of the first set of values corresponds to a midlife persistence of a respective one of the sets of birth and death coordinates for the plurality of topological features; and determining a second set of values, wherein each of the second set of values corresponds to a lifespan persistence of a respective one of the sets of birth and death coordinates for the plurality of topological features.

[0095] (6) The method of aspect 5, wherein determining the persistence statistics vector comprises: determining a first persistence statistics vector containing a mean, a standard, a skewness, a kurtosis, and/or an entropy of the first set of values; and determining a second

persistence statistics vector containing a mean, a standard, a skewness, a kurtosis, and/or an entropy of the second set of values.

[0096] (7) The method of aspect 6, wherein determining the persistence curve vector is performed based on the second set of values.

[0097] (8) The method of any of aspects 1-7, further comprising segmenting the 2D image before computing the sets of values for the plurality of topological features.

[0098] (9) The method of any of aspects 1-7, further comprising performing a smoothing operation on the 2D image before computing the sets of values for the plurality of topological features.

[0099] (10) The method of aspect 9, wherein the smoothing operation comprises a morphological closing operation.

[0100] (11) The method of any of aspects 1-10, wherein determining whether the 2D image is indicative of a defect in the manufactured component comprises using a trained statistical model and the digital feature as an input to the trained statistical model.

[0101] (12) The method of aspect 11, further comprising: determining a local binary pattern based on the 2D image; and using the trained statistical model and the digital feature comprises additionally using the local binary pattern as an input to the trained statistical model.

[0102] (13) The method of aspect 12, further comprising: determining a local binary pattern vector based on a histogram of the local binary pattern; wherein using the local binary pattern as the input to the trained statistical model comprises using the local binary pattern vector as the input to the trained statistical model.

[0103] (14) The method of any of aspects of 11-13, wherein the trained statistical model comprises an XGBoost classifier.

[0104] (15) The method of any of aspects 1-14, wherein receiving the 2D image of the manufactured component comprises capturing the 2D image of the manufactured component while the manufactured component is moving along a conveyor belt.

[0105] (16) The method of any of aspects 1-15, wherein: the manufactured component is a syringe; and capturing the 2D image of the manufactured component comprises capturing the 2D image from a top of the manufactured component.

[0106] (17) The method of any of aspects 1-16, wherein the defect comprises a crack, a chip, or both.

[0107] (18) The method of aspect 3, wherein computing the sets of birth and death coordinates comprises: varying the filtration threshold; and comparing intensities of the 2D image with the filtration threshold to identify the sets of birth and death coordinates.

[0108] (19) The method of aspect 18, wherein: each topological feature of the plurality of topological features comprises a shape comprising an isolated block or a block in a first color completely enclosed by a region in a second color, wherein each pixel in the shape is below the filtration threshold; and the birth coordinate for each topological feature corresponds to a first level of the filtration threshold at which said topological feature first appears as the filtration threshold is increased.

[0109] (20) The method of aspect 19, wherein the death coordinate for each topological feature specifies a second level for the filtration threshold at which said topological feature first disappears as the filtration threshold is increased.

[0110] (21) A system for detecting whether a manufactured component has a defect, the system comprising: an image capturing device configured to capture an image of the manufactured component; and at least one processor to perform the method of any of aspects 1-20.

[0111] (22) A non-transitory computer-readable media storing instructions that, when executed by at least one processor, cause the at least one processor to perform the method of any of aspects 1-20.

CLAIMS

1. A system for detecting whether a manufactured component has a defect, the system comprising:

an image capturing device configured to capture a two-dimensional (2D) image of the manufactured component;

at least one processor configured to execute computer-readable instruction to:

receive the 2D image of the manufactured component;

compute a respective set of values indicating a life for each topological feature of a plurality of topological features in the 2D image;

determine a digital feature based on the computed sets of values for the plurality of topological features; and

determine, based on the digital feature, whether the 2D image is indicative of a defect in the manufactured component.

2. The system of claim 1, wherein each topological feature of the plurality of topological features comprises a shape comprising an isolated block or a block in a first color completely enclosed by a region in a second color, wherein each pixel in the shape is below a filtration threshold.
3. The system of claim 1, wherein the set of values for each topological feature of the plurality of topological features in the 2D image include a set of birth and death coordinates each respectively representing a birth and a death of the topological feature.

4. The system of claim 3, wherein determining the digital feature comprises:

determining a persistence statistics vector, a persistence curve vector, or a combination thereof, based on the computed sets of birth and death coordinates for the plurality of topological features.

5. The system of claim 4, wherein the at least one processor is further configured to execute the computer-readable instructions to:

determine a first set of values, wherein each of the first set of values corresponds to a midlife persistence of a respective one of the sets of birth and death coordinates for the plurality of topological features; and

determine a second set of values, wherein each of the second set of values corresponds to a lifespan persistence of a respective one of the sets of birth and death coordinates for the plurality of topological features.

6. The system of claim 5, wherein determining the persistence statistics vector comprises:

determining a first persistence statistics vector containing a mean, a standard, a skewness, a kurtosis, and/or an entropy of the first set of values; and

determining a second persistence statistics vector containing a mean, a standard, a skewness, a kurtosis, and/or an entropy of the second set of values.

7. The system of claim 6, wherein determining the persistence curve vector is performed based on the second set of values.
8. The system of any of claims 1-7, wherein the at least one processor is further configured to execute the computer-readable instructions to segment the 2D image before computing the sets of values for the plurality of topological features.
9. The system of any of claims 1-7, wherein the at least one processor is further configured to perform a smoothing operation on the 2D image before computing the sets of values for the plurality of topological features.
10. The system of claim 9, wherein the smoothing operation comprises a morphological closing operation.
11. The system of any of claims 1-10, wherein determining whether the 2D image is indicative of a defect in the manufactured component comprises using a trained statistical model and the digital feature as an input to the trained statistical model.

12. The system of claim 11, wherein the at least one processor is further configured to execute the computer-readable instructions to:

determine a local binary pattern based on the 2D image; and

use the trained statistical model and the digital feature comprises additionally using the local binary pattern as an input to the trained statistical model.

13. The system of claim 12, wherein the at least one processor is further configured to execute the computer-readable instructions to:

determine a local binary pattern vector based on a histogram of the local binary pattern;

wherein using the local binary pattern as the input to the trained statistical model comprises using the local binary pattern vector as the input to the trained statistical model.

14. The system of any of claims of 11-13, wherein the trained statistical model comprises an XGBoost classifier.

15. The system of any of claims 1-14, wherein the image capturing device is configured to capture the 2D image of the manufactured component while the manufactured component is moving along a conveyor belt.

16. The system of any of claims 1-15, wherein:

the manufactured component is a syringe; and

capturing the 2D image of the manufactured component comprises capturing the 2D image from a top of the manufactured component.

17. The system of any of claims 1-16, wherein the defect comprises a crack, a chip, or both.

18. The system of claim 3, wherein computing the sets of birth and death coordinates comprises:

varying a filtration threshold; and

comparing intensities of the 2D image with the filtration threshold to identify the sets of birth and death coordinates.

19. The system of claim 18, wherein:

each topological feature of the plurality of topological features comprises a shape comprising an isolated block or a block in a first color completely enclosed by a region in a second color, wherein each pixel in the shape is below the filtration threshold; and

the birth coordinate for each topological feature corresponds to a first level of the filtration threshold at which said topological feature first appears as the filtration threshold is increased.

20. The system of claim 19, wherein the death coordinate for each topological feature specifies a second level for the filtration threshold at which said topological feature first disappears as the filtration threshold is increased.

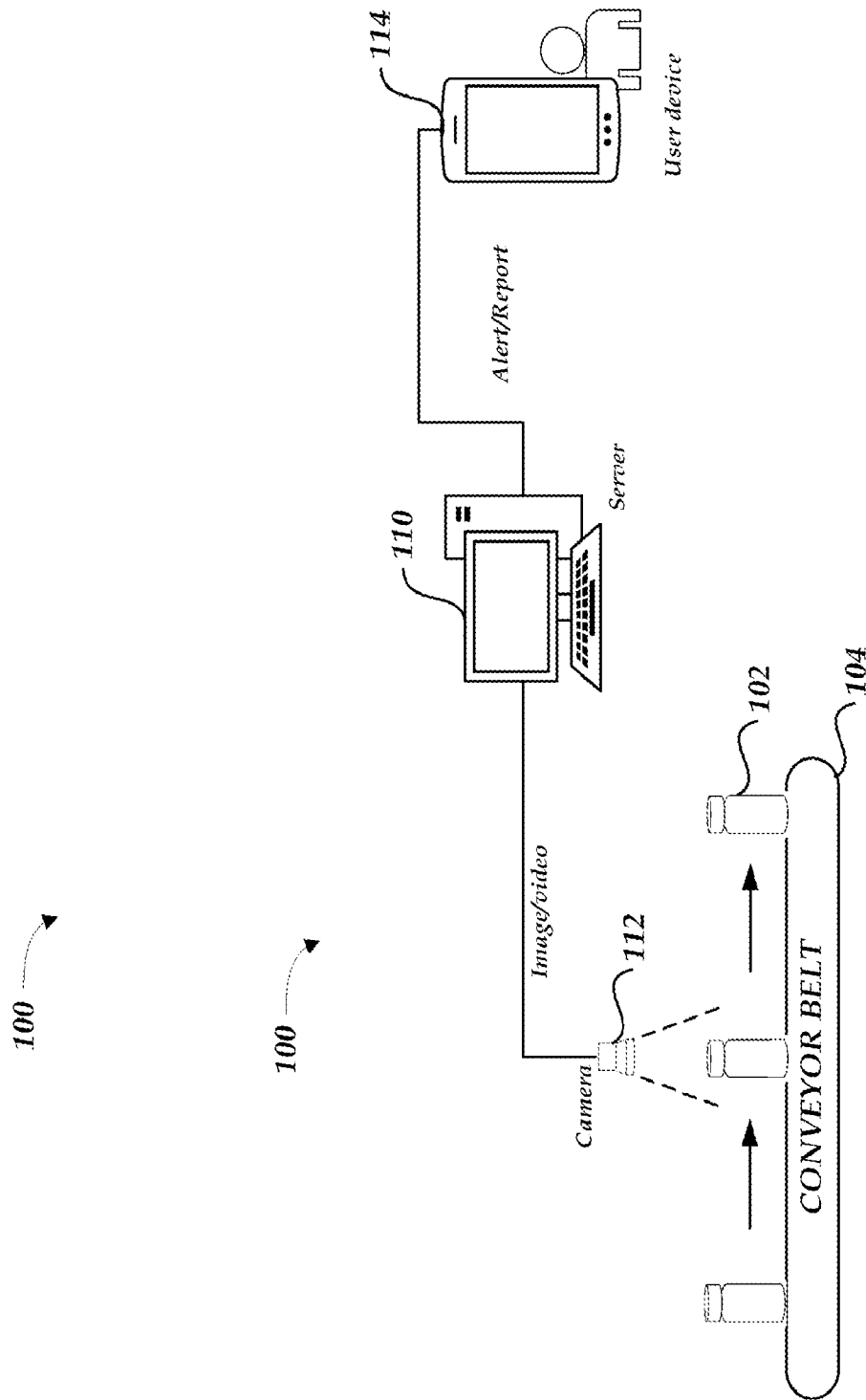


FIG. 1

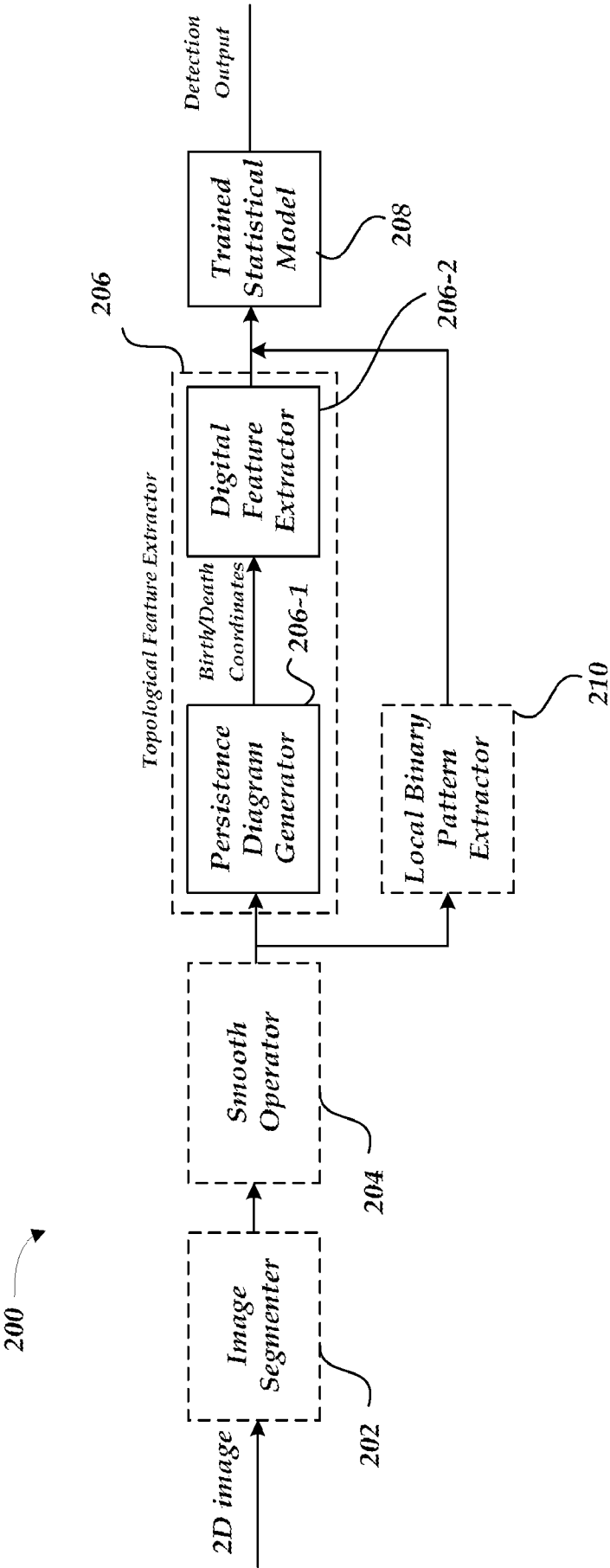
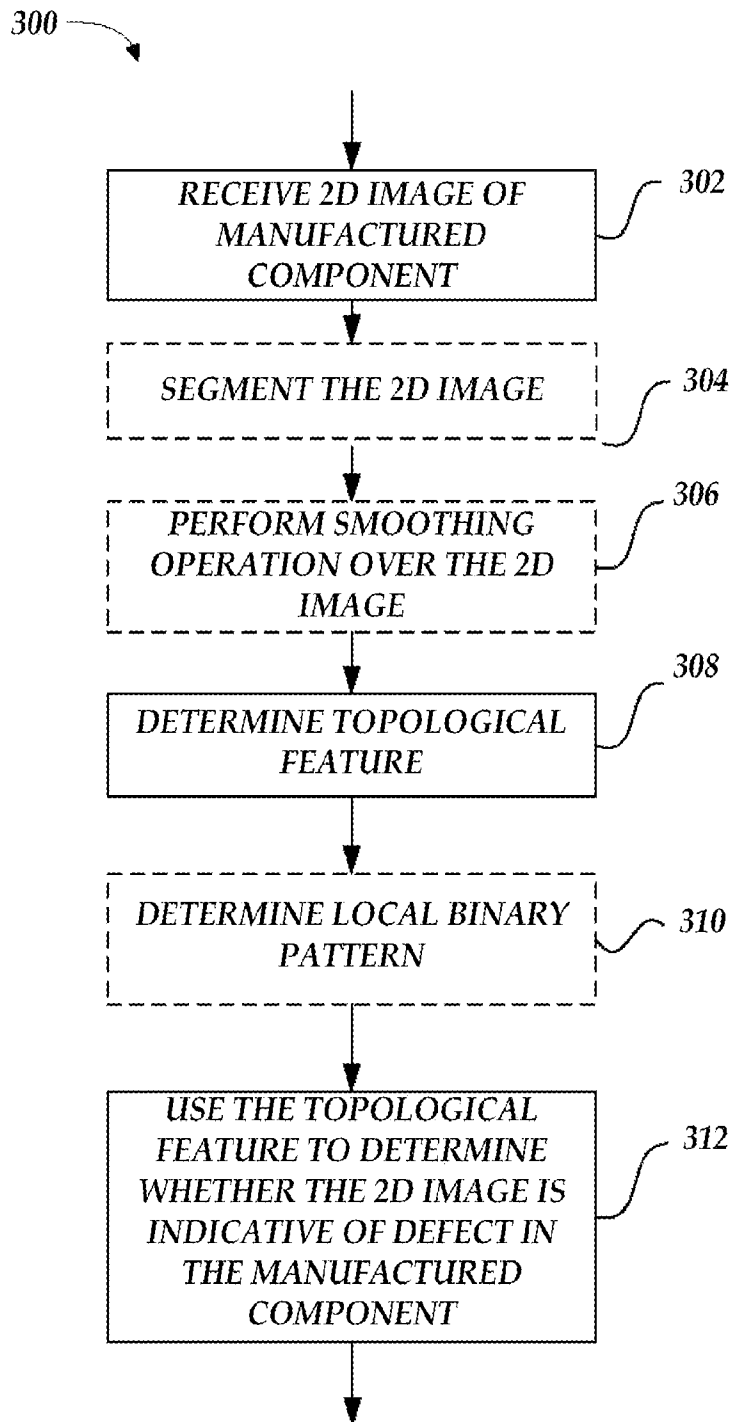
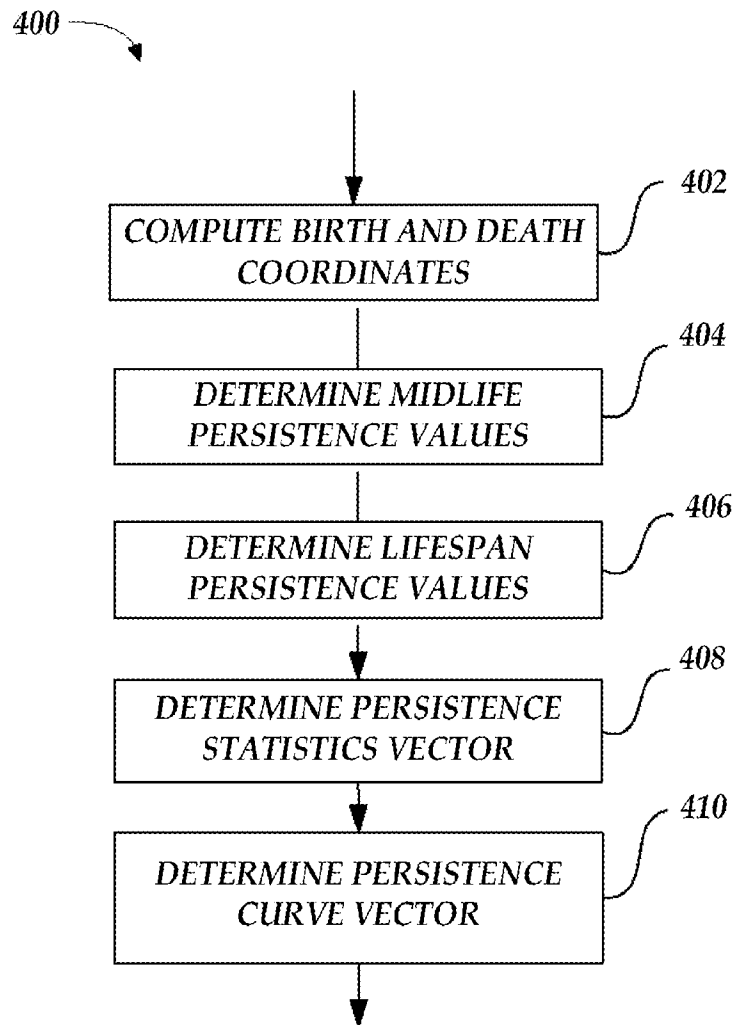
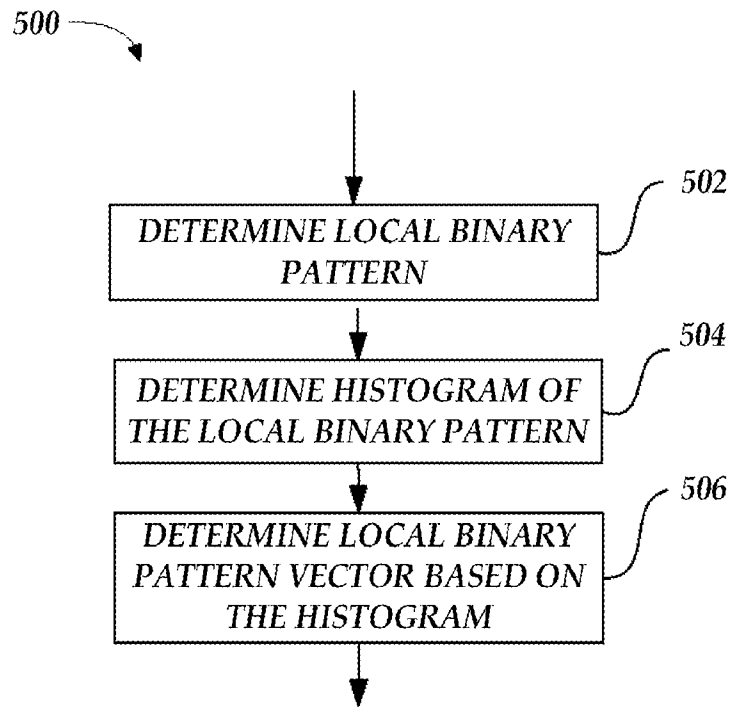
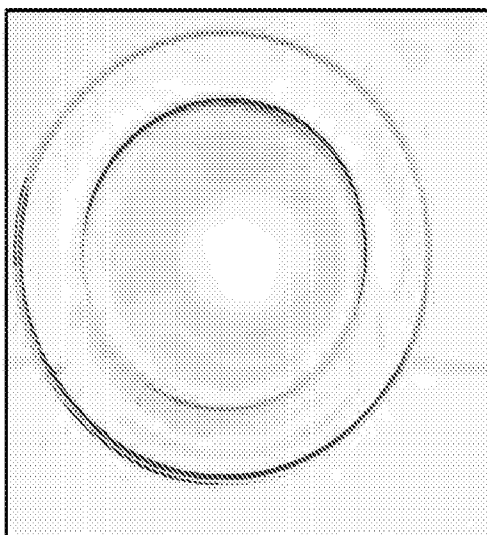
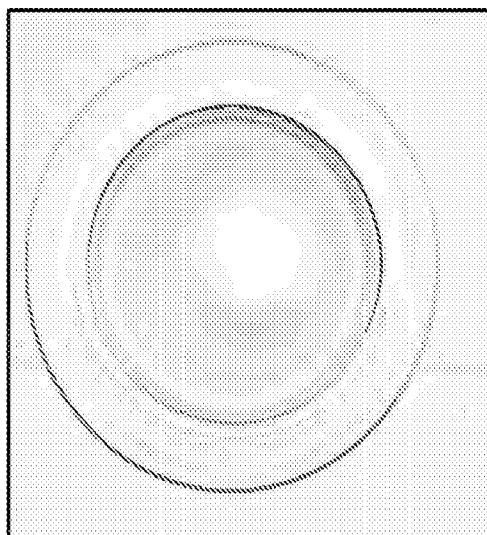
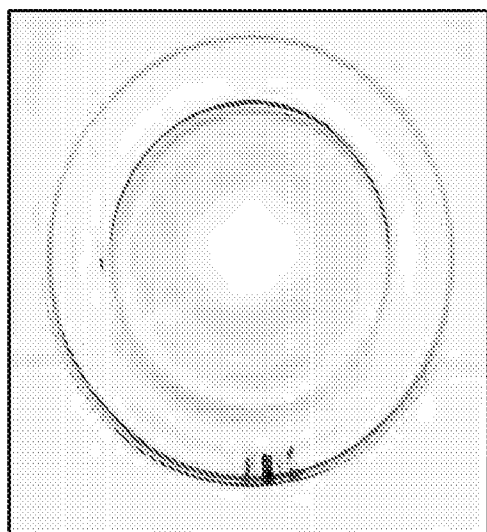
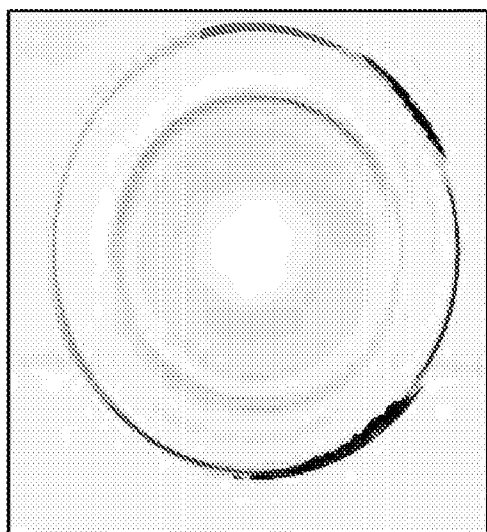


FIG. 2

**FIG. 3**

**FIG. 4**

***FIG. 5***

***FIG. 6A******FIG. 6B******FIG. 6C******FIG. 6D***

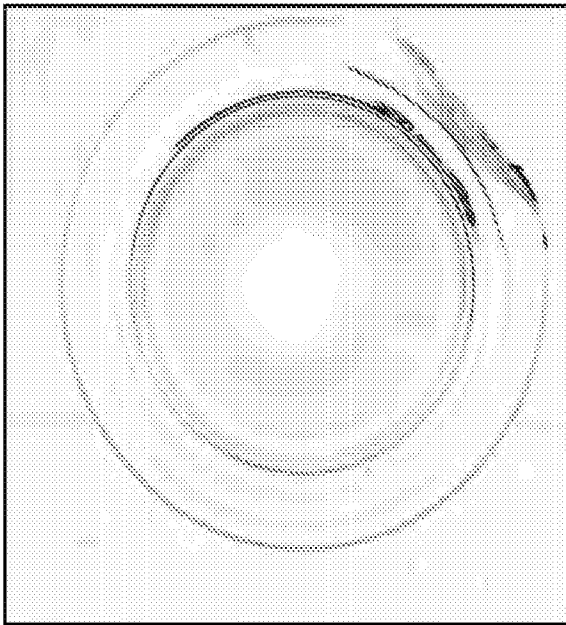


FIG. 6E

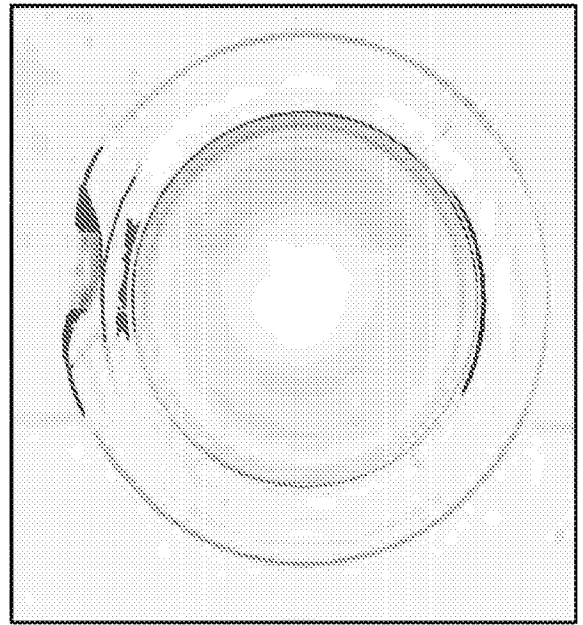


FIG. 6F

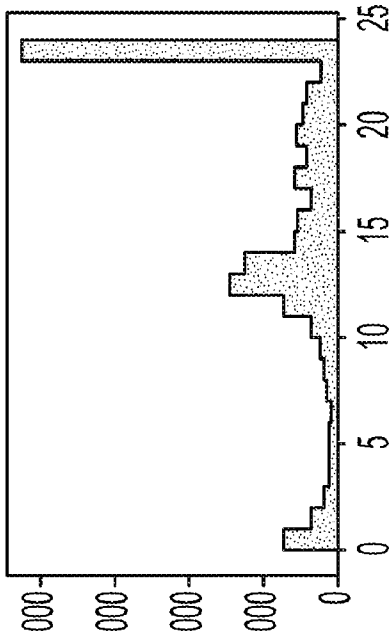


FIG. 7C

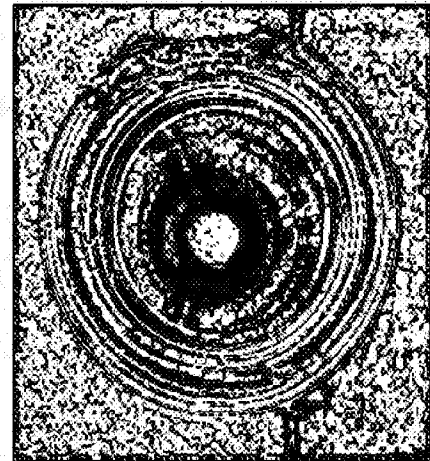


FIG. 7B

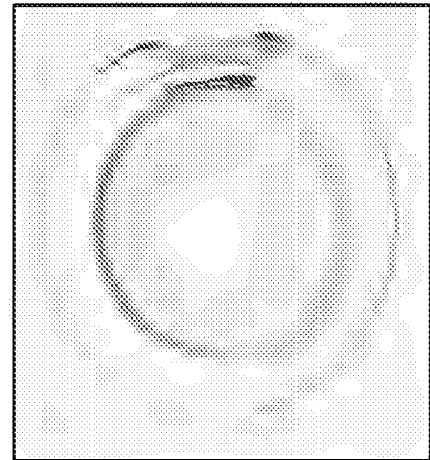


FIG. 7A

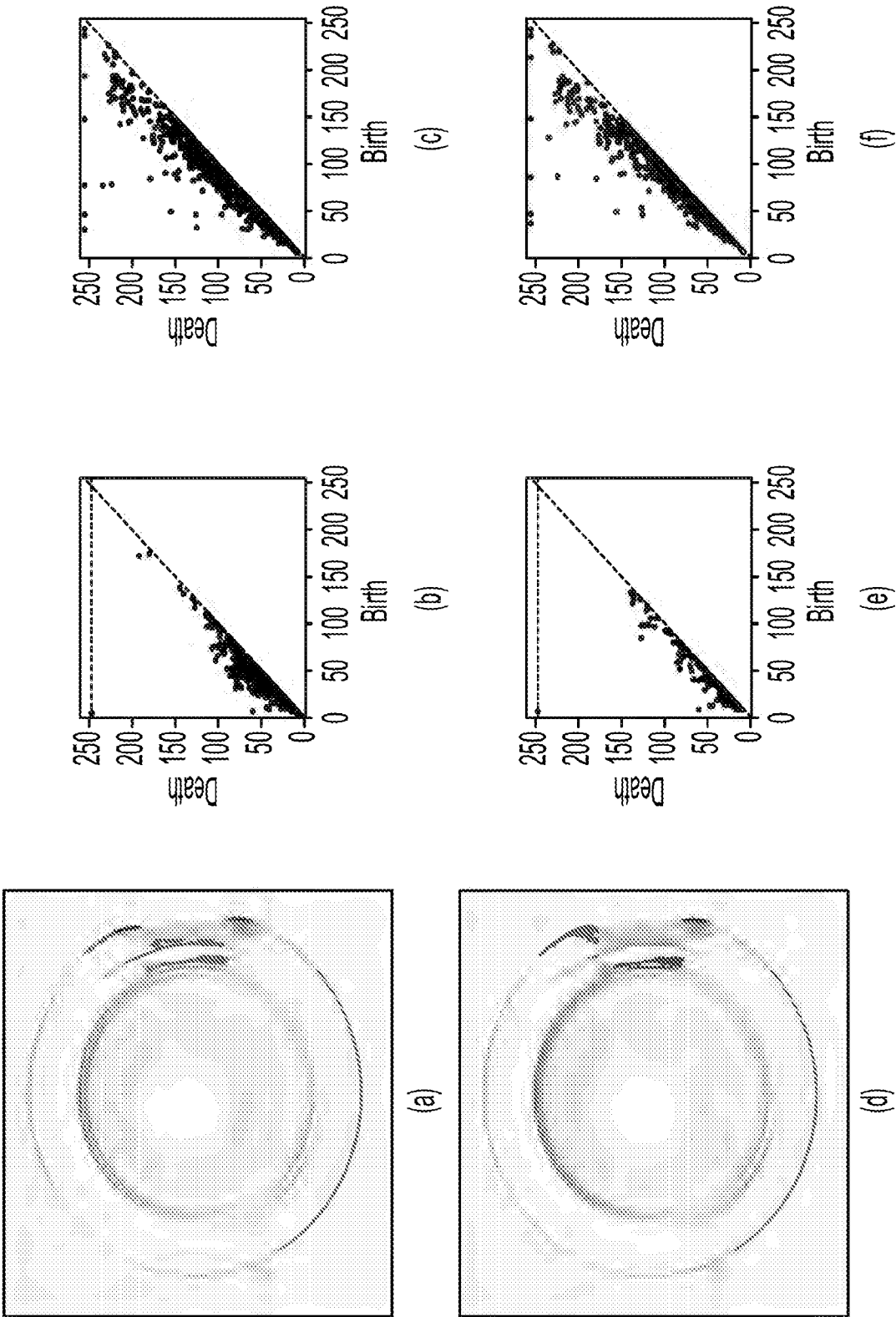


FIG. 8

1	3	2
1	10	2
1	3	2

FIG. 9A

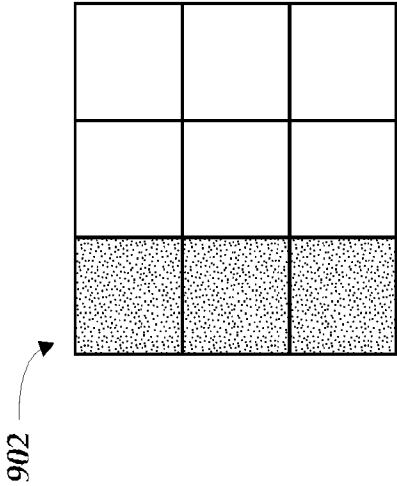


FIG. 9B

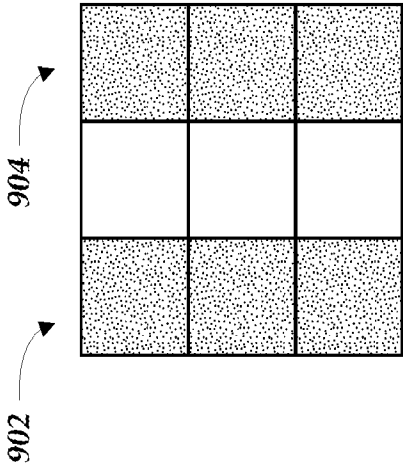


FIG. 9C

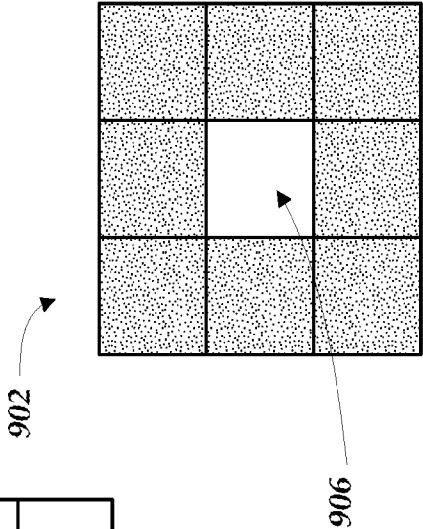


FIG. 9D

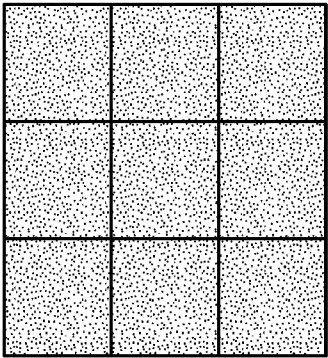


FIG. 9E

1000

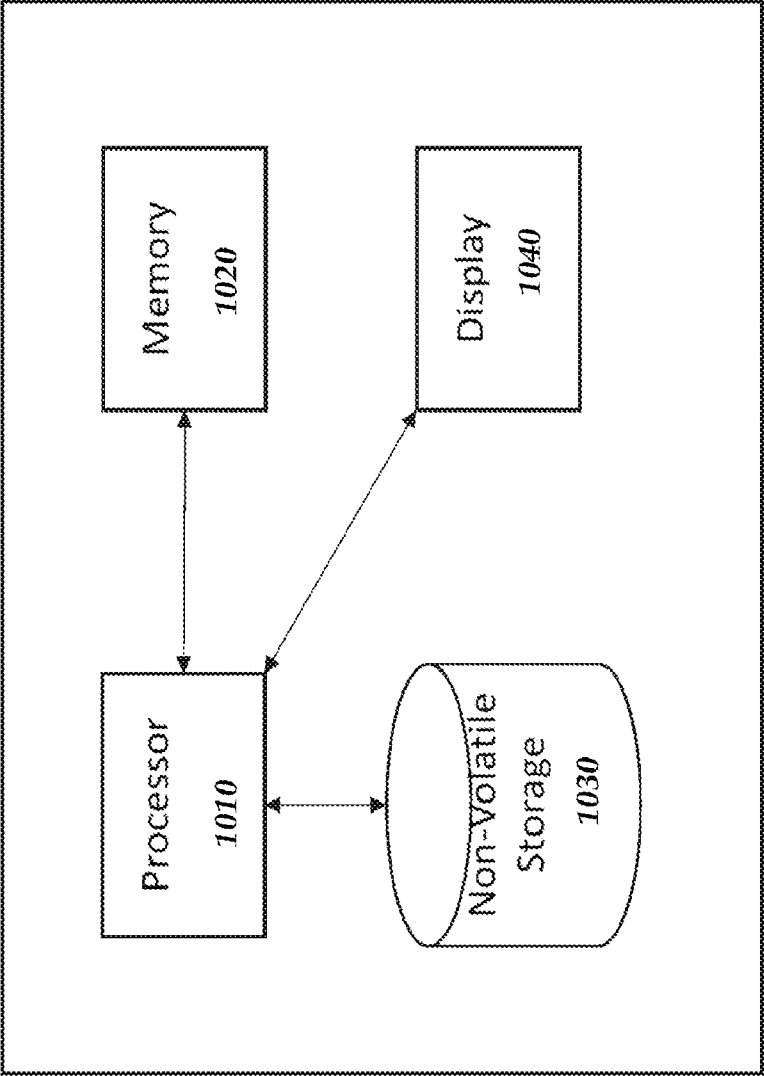


FIG. 10

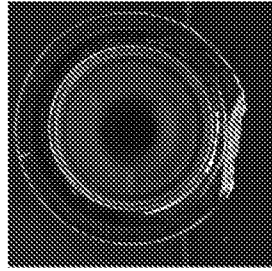


FIG. 10B

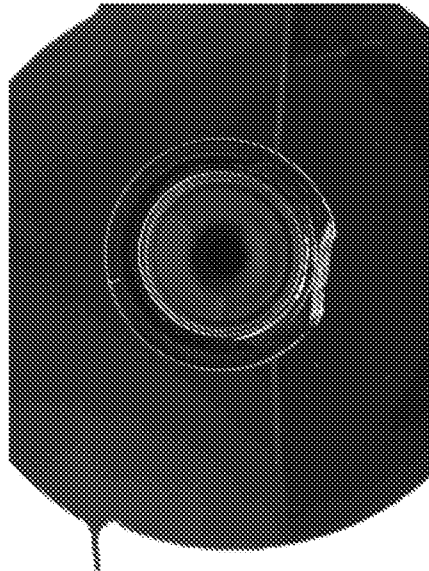


FIG. 10A

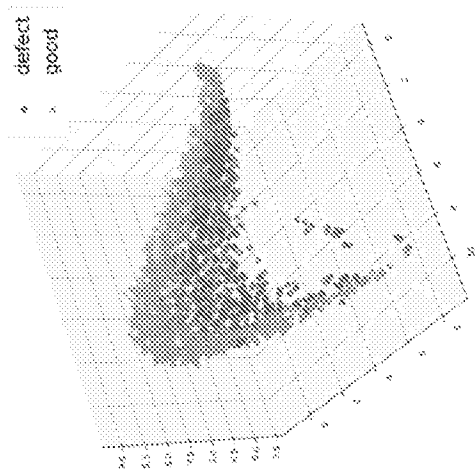


FIG. 11C

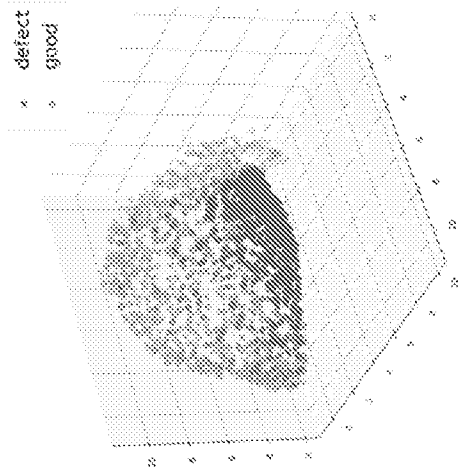


FIG. 11B

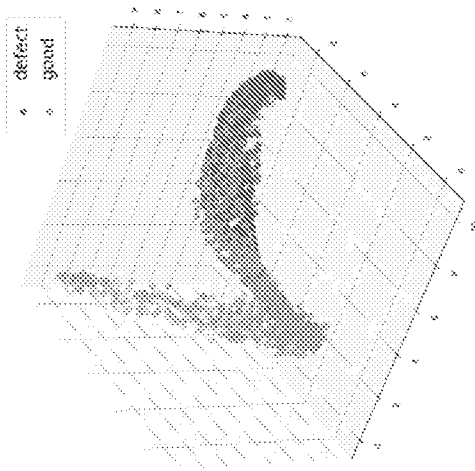


FIG. 11A

1000

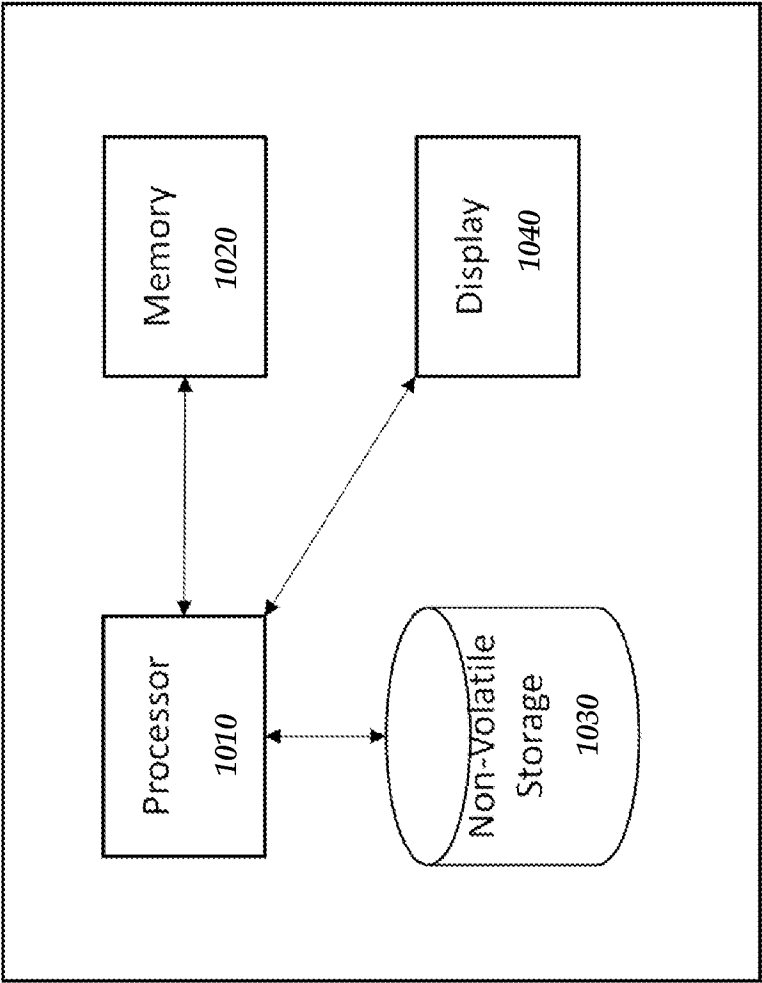


FIG. 12

INTERNATIONAL SEARCH REPORT

International application No
PCT/US2024/033334

A. CLASSIFICATION OF SUBJECT MATTER INV. G06T7/00 ADD.		
According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols) G06T		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used) EPO-Internal		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	MAX M CHUMLEY ET AL: "Pattern Characterization Using Topological Data Analysis: Application to Piezo Vibration Striking Treatment", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 12 October 2022 (2022-10-12), XP091342295,	1-4,8, 18-20
Y	Abstract; Sections 2, 3, 5; Figures 1, 4, 11 ----- -/-	5-7,9-17
☒ Further documents are listed in the continuation of Box C. ☒ See patent family annex.		
* Special categories of cited documents : "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance;; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance;; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search 3 October 2024		Date of mailing of the international search report 15/10/2024
Name and mailing address of the ISA/ European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Fax: (+31-70) 340-3016		Authorized officer Thean, Andrew

INTERNATIONAL SEARCH REPORT

International application No

PCT/US2024/033334

C(Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	CHUNG YU-MIN ET AL: "Topological approaches to skin disease image analysis", 2018 IEEE INTERNATIONAL CONFERENCE ON BIG DATA (BIG DATA), IEEE, 10 December 2018 (2018-12-10), pages 100-105, XP033508159, DOI: 10.1109/BIGDATA.2018.8622175 [retrieved on 2019-01-22] Section IV -----	5-7, 11-17
Y	ANDREW M THOMAS ET AL: "Feature detection and hypothesis testing for extremely noisy nanoparticle images using topological data analysis", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 18 January 2023 (2023-01-18), XP091415206, Section 3 -----	9,10
Y	CN 114 972 252 A (XCMG HANYUN TECH CO LTD) 30 August 2022 (2022-08-30) Figure 1, step 3; paragraph 2 -----	14,15
Y	US 2022/398715 A1 (CHAVALI NEELIMA [US] ET AL) 15 December 2022 (2022-12-15) Paragraph 0019 -----	16,17

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No
PCT/US2024/033334

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
CN 114972252	A	30-08-2022	NONE

US 2022398715	A1	15-12-2022	AU 2020378062 A1 07-04-2022
			BR 112022008676 A2 19-07-2022
			CA 3153701 A1 14-05-2021
			CL 2022001166 A1 10-02-2023
			CN 114631125 A 14-06-2022
			EP 4055559 A1 14-09-2022
			IL 291773 A 01-06-2022
			JP 2022553572 A 23-12-2022
			KR 20220090513 A 29-06-2022
			US 2022398715 A1 15-12-2022
			WO 2021092297 A1 14-05-2021
