

A Novel Approach for Defect Detection on Vessel Structures using Saliency-related Features

Francisco Bonnin-Pascual^{a,*}, Alberto Ortiz^a

^a*Department of Mathematics and Computer Science, University of the Balearic Islands, Cra. Valldemossa km 7.5, 07122 Palma de Mallorca, Spain*

Abstract

Seagoing vessels have to undergo regular visual inspections in order to detect defects such as cracks and corrosion before they result into catastrophic consequences. These inspections are currently performed manually by ship surveyors at a great cost, so that any level of assistance during the inspection process by means of e.g. a fleet of robots capable of defect detection would significantly decrease the inspection cost. In this paper, we describe a novel framework for visually detecting the aforementioned defects. This framework is generic and flexible in the sense that it can be easily configured to compute the features that perform better for the inspection at hand. Making use of this framework and inspired by the idea of conspicuity, this work considers contrast and symmetry as features for detecting defects and shows their usefulness for the case of vessels. Three different combination operators are additionally tested in order to merge the information provided by these features and improve the detection performance. Experimental results for different configurations of the detection framework show better classifi-

*Corresponding author: Tel.: +34-971-172-565;
Email address: xisco.bonnin@uib.es (Francisco Bonnin-Pascual)

cation rates than state of the art methods and prove its usability for images collected by a micro-aerial robotic platform intended for visual inspection.

Keywords:

Defect detection, Vessel inspection, Corrosion, Cracks, Saliency,
Micro-Aerial Vehicle

1. Introduction

Vessels are nowadays one of the most cost effective ways to transport goods around the world. Despite the efforts to avoid maritime accidents and wreckages, these still occur, and, from time to time, have catastrophic consequences in environmental, human and/or economic terms. Structural failures are the main cause of these accidents and, as such, Classification Societies impose extensive inspection schemes in order to ensure the structural integrity of vessels.

An important part of the vessel maintenance has to do with the visual inspection of the internal and external parts of the vessel hull. They can be affected by different kinds of defects typical of steel surfaces and structures, such as cracks and corrosion. These defects are indicators of the state of the metallic surface and, as such, an early detection prevents the structure from buckling and/or fracturing.

To carry out this task, the vessel has to be emptied and situated in a dockyard where scaffoldings are installed to allow the human inspectors to access the highest parts of the vessel structure (higher than 30 m in some cases). Taking into account the huge dimensions of some vessels, this process can mean the visual assessment of more than 600,000 m² of steel. Besides,

20 the surveys are on many occasions performed in hazardous environments for
21 which the access is usually difficult and the operational conditions turn out
22 to be sometimes extreme for human operation. Moreover, total expenses
23 involved by the infrastructure needed for close-up inspection of the hull can
24 reach up to one million dollars for certain sorts of vessels (e.g. Ultra Large
25 Crude Carriers). Therefore, it is clear that any level of automation of the
26 inspection process that can lead to a reduction of the inspection time, a
27 reduction of the financial costs, and/or an increase in the safety of the oper-
28 ation, is fully justified.

29 The EU-funded projects MINOAS (finished in 2012) and INCASS have
30 among their goals the development of robotic platforms to automate as much
31 as possible vessels' inspection processes (Eich et al., 2014). One of these
32 robots is a micro-aerial vehicle fitted with cameras, which is in charge of
33 collecting images that can provide the surveyor with a global overview of the
34 different surfaces and structures of the vessel (Bonnin-Pascual et al., 2015).
35 These images are intended to be processed afterwards to autonomously detect
36 the defective areas.

37 Previous approaches on vision-based defect detection can be roughly clas-
38 sified into two big categories. On the one hand, there are lots of contributions
39 on industrial inspection and quality control; that is to say, algorithms that
40 are in charge of checking whether the products that result from an industrial
41 manufacturing process are in good condition. These methods assume a more
42 or less confined environment where the product to be inspected is always
43 situated in a similar position, while lighting conditions are controlled as well.
44 Most of these techniques are collected in Chin and Harlow (1982); Newman

45 (1995); Malamas et al. (2003); Xie (2008).

46 On the other hand, several other contributions focus on visual inspection
47 techniques to ensure the integrity of elements or structures that have been
48 subjected to some kind of effort or stress. These methods are typically in-
49 cluded in periodical surveys to assess the need of maintenance operations. In
50 this group, which include vessel hull inspection, we can find algorithms for
51 crack detection on concrete surfaces (Yamaguchi and Hashimoto, 2010), de-
52 fect detection on bridge structures (Jahanshahi et al., 2009), aircraft surface
53 inspection (Siegel and Gunatilake, 1998; Mumtaz et al., 2010), etc.

54 The majority of the algorithms from both categories have been devised
55 for the detection of a specific defect on a particular material or surface, while
56 much less methods deal with unspecified defects on general surfaces. The
57 short distance from which the images must be captured is another point in
58 common among the majority of the algorithms. Furthermore, to provide
59 good results, most of them require from a learning and/or parameter-tuning
60 stages.

61 Special mention is made here to recent solutions based on Convolutional
62 Neural Networks (CNNs), adopting latest deep learning training approaches.
63 These techniques are widely used nowadays in many computer vision appli-
64 cations due to its high capacity of learning and their good performance in
65 non-easy classification problems. By way of example, Oullette et al. (2004)
66 and Zhang et al. (2016) describe methods based on CNNs for the detection of
67 cracks, while the approach presented by Petricca et al. (2016) focuses on the
68 detection of corrosion. As mentioned before, these machine learning tech-
69 niques require from a previous training stage, which, in this case, involves a

70 very large dataset.

71 Regarding defect detection over vessel structures, just a few contribu-
72 tions can be found. For example, Ozog and Eustice (2015) present a method
73 to identify structural anomalies over visual reconstructions of underwater
74 ship hulls. Restricting to those contributions which just use visual sensors,
75 Bonnin-Pascual (2010) and Bonnin-Pascual and Ortiz (2014b) present de-
76 tectors of cracks and corrosion for vessel structures. These algorithms do
77 not need close-up images of the inspected surfaces to provide good results
78 but their drawback is again that they require a previous training stage (e.g.
79 to learn which is the color that corrosion usually presents) or tuning their
80 working parameters (e.g. to know how thin and elongated must be a dark
81 collection of pixels to be considered a crack), whose value is typically related
82 with the distance from which the images have been collected.

83 To the best of our knowledge, only one method has been published for
84 generic defect detection in vessel structure images (Bonnin-Pascual and Or-
85 tiz, 2014a). This approach makes use of a Bayesian framework to compute
86 the probability of every pixel to correspond to some kind of defective sit-
87 uation. This probability is based on the information learned in a previous
88 training stage.

89 This paper presents a novel approach for automatic detection of defects
90 in images taken from the vessel structures. Unlike previous works, the pre-
91 sented approach does not require from tuning a large set of parameters nor
92 performing a previous training stage. A framework is proposed as a generic
93 classifier that can be configured to make use of different features, poten-
94 tially leading to different defect detectors each. Furthermore, the framework

95 foresees the combination of the respective feature responses in order to en-
96 hance the overall output quality. The conspicuousness of defects in general,
97 together with the kind of defects that can be expected in metallic surfaces
98 (i.e. cracks and corrosion) and the image capture conditions, have guided
99 the feature selection process.

100 The rest of the paper is organized as follows: Section 2 describes the
101 generic flexible defect detection framework; Section 3 explains how this frame-
102 work particularizes for defect detection in vessel structures, considering con-
103 trast (3.1), symmetry (3.2) and three alternative combinations among them
104 (3.3); Section 4 discusses on the results of some experiments; and Section 5
105 concludes the paper.

106 **2. A Flexible Framework for Defect Detection**

107 The importance of feature selection during the design of any classifier is
108 discussed in Theodoridis and Koutroumbas (2006). In particular, the follow-
109 ing questions must be answered: (1) which features are the best for a suitable
110 classification, (2) how many features are necessary, and (3) how should these
111 be combined to implement the best classifier.

112 Taking that into account, we oriented the design of our defect detector
113 towards a flexible framework which allows an easy integration of different
114 features and their combinations. To attain this level of flexibility, we consid-
115 ered that the framework must cover the following aspects: (1) it should allow
116 computing one or more features that are potentially useful to discriminate
117 between defective and non-defective situations; (2) final features response
118 should not depend on scale; (3) one or more combination operators should

119 be available to merge the information provided by the computed features
120 and try to find the combination (if any) that improves the classification per-
121 formance; and (4), related to the previous point, one or more normalization
122 operators should be available to adapt the different features responses to a
123 certain range, in order to ensure a proper combination.

124 This generic framework has been organized as a modular pipeline which
125 involves different stages that can be configured (or even removed) depend-
126 ing on our needs, so that different configurations result into different defect
127 detectors (see Fig. 1). Within the framework, each feature is computed as a
128 different thread, while the final detection output results from the combination
129 of the information supplied by all of them.

130 In more detail, the framework consists of the following stages:

- 131 • *Pre-feature computation.* The first stage prepares the input image to
132 provide the information necessary to compute all features. From an
133 input color image one can obtain, for example, the gray-scale (or inten-
134 sity) image, the red channel image, the saturation image (from HSV
135 color space), etc. Each one of these images is called a *pre-feature map*.
- 136 • *Scale-space generation.* This stage scales the pre-feature maps using a
137 range of scale factors to obtain a collection of multiple-scale represen-
138 tations, also known as pyramids. The computation of each pyramid
139 level can include filtering the input map using a specific kind of filter.
140 One can compute, for example, a Gaussian pyramid which progressively
141 low-pass filters and sub-samples the pre-feature map, an oriented Gabor
142 pyramid for a preferred orientation θ , a simple sub-sampling pyramid
143 computed without any filtering, etc.

- 144 • *Feature computation.* This is the core stage within the pipeline. Each
145 instance of this stage is in charge of computing the value for a given
146 feature for all the pixels of the input image. Since this can be fed with
147 one or more multi-scale pyramids, a feature can be computed combining
148 the information provided at different scales. Every output of this stage
149 is called a *feature map*.
- 150 • *Normalization.* This step normalizes the different feature maps to the
151 same range of values to enable their combination.
- 152 • *Combination.* This is the last stage of the pipeline. It is in charge of
153 combining the normalized feature maps in order to obtain a single map,
154 which is called the *defect map*. The mean and the median operators
155 are some examples of simple combination operators. Unary operators
156 such as unary minus or thresholding can also be considered.

157 As indicated in Fig. 1, the generic framework allows computing more com-
158 plex features by means of concatenating multiple instances of normalization
159 and combination stages.

160 The output of the framework is the defect map, which consists in a single-
161 channel map where defective areas are supposed to be labelled with higher
162 values.

163 **3. Detecting Defects on Vessel Structures**

164 Vessel structures consist of large surfaces that usually present a regular
165 texture. When these surfaces are inspected from a certain distance, a defect
166 appears as a discontinuity that alters the regularity of the texture. Based

167 on that, texture-related features seem to be a good option to differentiate
168 between defective and non-defective areas.

169 Furthermore, defects can also be considered as rare phenomena that may
170 appear on such regular surfaces. Since they are rare, defects potentially
171 attract the visual attention of the surveyor during a visual inspection process.
172 Following these ideas, we propose to use texture-based features typically used
173 in cognitive models to predict human eye fixations.

174 Among them, we focus on those which can be evaluated through a saliency
175 map. A saliency map consists in a topographic map that represents the
176 conspicuousness of the different areas of the input image (Koch and Ullman,
177 1985). This is typically shown as a gray-scale image where locations with
178 higher conspicuity values are closer to white and less salient areas are closer
179 to black. Notice that this representation fits with our definition of defect
180 map.

181 Taking all these considerations into account, contrast and symmetry have
182 been selected as the features for detecting defects on vessel structures. The
183 following sections detail further about the motivations that led us to consider
184 these features as well as describe how the defect detector makes use of them.

185 *3.1. The Contrast-based Defect Detector*

186 As indicated in Borji and Itti (2013), three features have been tradition-
187 ally used in computational models of attention: intensity, color and orien-
188 tation. The sudden variation of some of these features, computed as a lo-
189 cal contrast, increases the conspicuousness of the area producing bottom-up
190 guidance (Wolfe, 2007).

191 The information resulting from the variation of these three features is

192 typically combined into a single contrast-based saliency map. See for example
 193 Avraham and Lindenbaum (2010); Borji et al. (2011); Li et al. (2010); Zhang
 194 et al. (2015).

195 We propose to use this local contrast (combining intensity, color and
 196 orientation) in a first attempt to locate coating breakdown/corrosion and
 197 cracks on vessel structures.

198 The generic framework described in section 2 is of application now to
 199 design the contrast-based defect detector. The model presented in Itti et al.
 200 (1998) has been used as source of inspiration to design the different stages of
 201 the pipeline. The previous work described for the first time a contrast-based
 202 model for saliency and has also inspired later authors (Borji and Itti, 2013).

203 Figure 3 details the contrast-based defect detector. As for its imple-
 204 mentation, each one of the stages of the generic pipeline (Fig. 1) has been
 205 particularized as follows:

- 206 • *Pre-feature computation.* Five pre-feature maps are computed from the
 207 red (r), green (g) and blue (b) channels of the input image:

$$I = \frac{r + g + b}{3}, \quad (1)$$

$$R = r - \frac{g + b}{2}, \quad (2)$$

$$G = g - \frac{r + b}{2}, \quad (3)$$

$$B = b - \frac{r + g}{2}, \quad (4)$$

$$Y = \frac{r + g}{2} - \frac{|r - g|}{2} - b, \quad (5)$$

212 where I is the intensity map, R is the red channel map, G is the green
 213 channel map, B is the blue channel map and Y is the yellow channel
 214 map. During the computation of these maps, negative values (if any)
 215 are set to zero.

- 216 • *Scale-space generation.* Nine pyramids are computed from the pre-
 217 feature maps. On the one hand, five Gaussian pyramids \widehat{I} , \widehat{R} , \widehat{G} , \widehat{B} and
 218 \widehat{Y} are computed by progressively low-pass filtering and sub-sampling
 219 the pre-feature maps (I , R , G , B and Y). On the other hand, four Ga-
 220 bor pyramids \widehat{O}_0 , \widehat{O}_{45} , \widehat{O}_{90} and \widehat{O}_{135} are computed filtering the images
 221 of the intensity pyramid \widehat{I} with oriented Gabor filters with orientations
 222 $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. All pyramids comprise seven scales, ranging
 223 from 1:1 (scale one) to 1:64 (scale seven).

- 224 • *Feature computation.* Three threads are executed in parallel to build
 225 three feature maps, respectively corresponding to the contrast level in
 226 intensity (\mathbf{I}), color (\mathbf{C}) and orientation (\mathbf{O}). This computation is per-
 227 formed as indicated in Itti et al. (1998). A first step computes center-
 228 surround differences between fine and coarse scales from the pyramids;
 229 that is, it computes the difference between each pixel of a fine (or cen-
 230 ter) scale c and its corresponding pixel in a coarse (or surrounding)
 231 scale s . Accordingly, preliminary maps $\mathbf{I}(c, s)$, $\mathbf{RG}(c, s)$, $\mathbf{BY}(c, s)$ and
 232 $\mathbf{O}(c, s, \theta)$ are created as follows:

$$\mathbf{I}(c, s) = |I(c) \ominus I(s)|, \quad (6)$$

$$\mathbf{RG}(c, s) = |R(c) - G(c) \ominus (G(s) - R(s))|, \quad (7)$$

234

$$\mathbf{BY}(c, s) = |(B(c) - Y(c)) \ominus (Y(s) - B(s))|, \quad (8)$$

235

$$\mathbf{O}(c, s, \theta) = |O(c, \theta) \ominus O(s, \theta)|, \quad (9)$$

236

237

238

239

240

241

where $|x|$ refers to the absolute value of x , \ominus is the across-scale subtraction operator (see Fig. 2), $\mathbf{I}(c, s)$ accounts for the intensity contrast, $\mathbf{RG}(c, s)$ accounts for red/green contrast, $\mathbf{BY}(c, s)$ accounts for blue/yellow contrast and $\mathbf{O}(c, s, \theta)$ accounts for the orientation contrast for a given orientation θ . In our implementation, the scales are defined as $c \in \{1, 2, 3\}$ and $s = c + \delta$, with $\delta \in \{3, 4\}$.

242

243

244

In a second step, the intermediate maps are combined into the following three feature maps by means of the across-scale addition operator \oplus (see Fig. 2 for details):

$$\mathbf{I} = \bigoplus_{c=1}^3 \bigoplus_{s=c+3}^{c+4} N(\mathbf{I}(c, s)), \quad (10)$$

$$\mathbf{C} = \bigoplus_{c=1}^3 \bigoplus_{s=c+3}^{c+4} (N(\mathbf{RG}(c, s)) + N(\mathbf{BY}(c, s))), \quad (11)$$

$$\mathbf{O} = \sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} N \left(\bigoplus_{c=1}^3 \bigoplus_{s=c+3}^{c+4} N(\mathbf{O}(c, s, \theta)) \right), \quad (12)$$

245

246

247

248

where $N(\cdot)$ is a normalization operator devised to promote high and isolated peaks. It adjusts the map to a fixed range $[0..M]$ and multiplies it by $(M - \bar{m})^2$, being \bar{m} the average of all local maxima that do not coincide with the global maximum.

249 By way of illustration, a diagram showing the entire feature computa-
250 tion for map \mathbf{I} can be found in Fig. 2.

- 251 • *Normalization.* The normalization operator $N(.)$ is used now to pro-
252 mote the highest and isolated peaks in the three feature maps, obtaining
253 $\bar{\mathbf{I}}$ for intensity, $\bar{\mathbf{C}}$ for color and $\bar{\mathbf{O}}$ for orientation.
- 254 • *Combination.* The final defect map is computed using a linear combi-
255 nation:

$$D_{con} = \frac{\bar{\mathbf{I}} + \bar{\mathbf{C}} + \bar{\mathbf{O}}}{3}, \quad (13)$$

256 so that any salient point in any of the feature maps appears in the final
257 defect map.

258 3.2. The Symmetry-based Defect Detector

259 A saliency model based on the Gestalt principle of symmetry was pre-
260 sented in Kootstra et al. (2008). In their paper, they discuss local symmetry
261 as a measure of saliency and investigate its role in visual attention. To
262 this end, they use three different symmetry operators (isotropic, radial and
263 color symmetry operators) and compare them with human eye tracking data.
264 Their results suggested that symmetry was a salient structural feature for
265 humans, as well as the suitability of their method for predicting human eye
266 fixations in complex photographic images, where symmetry is not so evident.

267 Furthermore, the authors use the saliency model by Itti et al. as a ref-
268 erence for comparison. Their results show that, on many occasions, their
269 symmetry operators outperformed the contrast-saliency model.

270 These are the reasons why, in this work, we decided to incorporate symme-
 271 try as a second feature for defect detection. Figure 4 shows our implemen-
 272 tation of the symmetry-based defect detector using the generic framework
 273 (Fig. 1), where each stage is particularized as follows:

- 274 • *Pre-feature computation.* It computes one intensity map as indicated
 275 in Eq. 1.
- 276 • *Scale-space generation.* This stage computes a simple sub-sampling
 277 pyramid with five scales, ranging from 1:1 (scale one) to 1:16 (scale
 278 five).
- 279 • *Feature computation.* The symmetry map is calculated for each level
 280 l of the pyramid, using the isotropic operator. We have chosen this
 281 operator because it is easier to compute and no significant improvement
 282 was observed when using the radial or color symmetry operators for
 283 predicting human eye fixations (Kootstra and Schomaker, 2009).

284 To obtain the final defect map based on symmetry, the five responses
 285 $\mathbf{M}(l)$ (one per pyramid level) are normalized using the normalization
 286 operator $N(\cdot)$ and finally added together across-scale into an scale 1:1
 287 map:

$$\mathbf{D}_{sym} = \bigoplus_{l=1}^5 N(\mathbf{M}(l)). \quad (14)$$

288 Normalization and combination stages are not employed for this case since
 289 symmetry is the only feature used.

290 *3.3. Combination of Contrast and Symmetry*

291 In order to deeper explore the possibilities of the selected features, the
 292 generic framework has been configured to combine the information that they
 293 convey in the following way:

- 294 • *Pre-feature computation.* Five pre-feature maps are computed as de-
 295 scribed for the contrast-based method.
- 296 • *Scale-space generation.* It generates ten pyramids, nine used for con-
 297 trast plus one used for symmetry, as detailed in, respectively, sections
 298 3.1 and 3.2.
- 299 • *Feature computation.* It consists of four threads, one for each channel of
 300 contrast (intensity, color and orientation) plus one for symmetry. They
 301 proceed as indicated in previous sections.
- 302 • *Normalization.* The normalization operator $N(\cdot)$ of section 3.1 is used
 303 in this stage to promote the areas from the feature maps that have been
 304 indicated as potentially defective by any of the features. Therefore,
 305 $\overline{\mathbf{D}}_{con}$ is obtained as the normalized version of the defect map based on
 306 contrast and $\overline{\mathbf{D}}_{sym}$ is the analogue for the case of symmetry.
- 307 • *Combination.* We initially propose two operators. The first one con-
 308 sists in a linear combination of the contrast and symmetry-based defect
 309 maps:

$$\mathbf{D}_{OR} = \frac{\overline{\mathbf{D}}_{con} + \overline{\mathbf{D}}_{sym}}{2}. \quad (15)$$

310 This combination allows any defective point in any of the maps to be
 311 promoted so that it stands out in the final defective map. From now on,

312 it will be referred to as the *OR* combination, since this operator labels
 313 as defective those areas that result defective in the contrast-based map
 314 ‘or’ in the symmetry-based map.

315 The second combination operator that we propose merges the contrast
 316 and symmetry-based defect maps so that defective regions in the re-
 317 sulting map are required to be simultaneously indicated as potentially
 318 defective in both maps:

$$D_{AND} = \overline{D_{con}} \times \overline{D_{sym}}, \quad (16)$$

319 implementing, in a certain sense, the *AND* operator, i.e. this operator
 320 only labels as defective those areas that are indicated in the contrast
 321 ‘and’ in the symmetry-based maps.

322 In addition to these combinations, a third version has been considered
 323 which intends to explore the contribution provided by the different
 324 contrast channels, i.e., intensity, color and orientation. The four feature
 325 maps (including the symmetry map) are fused using a modified version
 326 of the OR combination:

$$D_{ORA} = \frac{\overline{I} + \overline{C} + \overline{O} + \overline{D_{sym}}}{4}, \quad (17)$$

327 which will be referred to as the *ORA* (Or-Alternative) combination,
 328 i.e. this operator labels as defective those areas that result defective
 329 in the intensity-contrast map ‘or’ in the color-contrast map ‘or’ in the
 330 orientation-contrast map ‘or’ in the symmetry map.

331 Figure 5 shows the set up of the normalization and combination stages
 332 for the three detectors which combine contrast and symmetry information.

333 4. Assessment of the Defect Detector

334 In this study, we have used a dataset comprising 73 images of vessel
335 structures including defective areas (cracks, coating breakdown and different
336 kinds of corrosion). The images have been collected at different distances
337 and under different lighting conditions. This dataset also includes the ground
338 truth consisting in binary images where defects are labelled in white (see Fig.
339 7:B), and it is available online (<http://dmi.uib.es/~xbonnin/resources>).

340 In this section, we report on the results of a number of experiments ori-
341 ented to determine the performance of the different defect detectors described
342 in the previous sections. In a first kind of experiment, we have assessed how
343 suitable are contrast and symmetry to differentiate between defective and
344 non-defective areas. To this end, the probability distribution of these two
345 features has been computed for the two classes, defective and non-defective
346 area. To estimate these PDFs, we have applied the Parzen windows method
347 (Theodoridis and Koutroumbas, 2006) to the histograms corresponding to
348 the combinations contrast/defect, symmetry/defect, contrast/non-defect and
349 symmetry/non-defect. The resulting PDFs are shown in Fig. 6. We can state
350 the following looking at those PDFs:

- 351 • non-defective pixels present low values of contrast and symmetry (below
352 10 for contrast and around 15 for symmetry), what confirms that non-
353 defective areas present an uniform-intensity texture;
- 354 • defective pixels tend to present higher values of both features (around
355 25), so that these features seem to be useful to differentiate between
356 defective and non-defective areas;

357 • contrast peaks are farther from one another than symmetry peaks,
358 what could indicate that contrast is more discriminative than symmetry
359 when describing the defective areas that appear in our dataset.

360 In a second kind of experiment, we evaluated the performance of the
361 proposed defect detectors. Figure 7 presents some examples of defect maps
362 provided for the different cases, namely, the contrast-based detector, the
363 symmetry-based detector and the three detectors which combine these two
364 features using, respectively, the OR, AND and ORA combination operators.

365 At first sight, it can be observed that all the defect detectors tend to label
366 as whitish the areas that are indicated as defective in the ground truth image.
367 This suggests that the different detectors can attain good classification rates.

368 In order to perform a quantitative evaluation, the True Positive Rate
369 (TPR), also known as recall, and the False Positive Rate (FPR), also known
370 as fall-out, have been computed for the five defect detectors. To this end,
371 the defect maps were thresholded for different values of a threshold τ to
372 obtain the corresponding ROC curves, which are presented in Fig. 8 [left].
373 Furthermore, the values for the Area Under the Curve (AUC) (Fawcett, 2006)
374 have been calculated for all the ROC curves, obtaining the values also shown
375 in Fig. 8 [left].

376 Comparing the different ROC curves and AUC values, some interesting
377 results can be stated: (1) the five defect detectors present good performances
378 during the classification task, with ROC curves relatively close to the (0,1)
379 corner (corresponding to the perfect classifier), and AUC values above 0.8;
380 (2) contrast performs better than symmetry for the dataset employed in this
381 study, what suggests that contrast provides more information to discriminate

382 between defective and non-defective areas in vessel structures; (3) the three
383 detectors which combine both contrast and symmetry information lead to
384 slightly better results than the version based only on contrast (i.e. symmetry
385 provides complementary information), being the ORA combination the one
386 which yields the highest AUC value.

387 Similarly, Precision-Recall (PR) curves are reported for the five defect
388 detectors. The precision indicates the proportion of positively classified sam-
389 ples (i.e. pixels classified as defective) which are actually positive. Informally
390 speaking, a high precision value means a low number of false positives, while
391 a high recall (TPR) value means a low number of false negatives. The PR
392 curves, which are provided in Fig. 8 [right], show that the combined detectors
393 attain higher precision values than the single-feature detectors.

394 In a third kind of experiment, the performances of the defect detectors
395 presented in this paper have been compared with the one attained by some
396 state of the art defect detectors. Each comparative assessment is performed
397 using ROC/PR curves, which are provided in separate plots to simplify their
398 interpretation. In a first experiment, we have compared with the WCCD
399 algorithm (Bonnin-Pascual, 2010). This algorithm was devised for corro-
400 sion detection in images taken from vessel structures. It consists in a cas-
401 cade classifier that combines texture (described as the energy of a gray-level
402 co-occurrence matrix downsampled to 32×32 intensity levels) and colour in-
403 formation, and which has proved to outperform other more complex weak-
404 classifier combinations, such as the ABCD algorithm (Bonnin-Pascual and
405 Ortiz, 2014b), which combines Laws' texture energy filters within an Ad-
406 aBoost framework. Notice that both WCCD and ABCD follow a supervised

407 classification scheme, so that they require from a previous training stage.

408 The WCCD algorithm has been slightly modified with regard to Bonnin-
409 Pascual (2010) to compute the energy for all the pixels of the image instead
410 of computing it only at the patch level (the same energy value was originally
411 used for all pixels of the 15×15 pixels patches), in order to obtain finer
412 classification results.

413 To perform the assessment, the original dataset was reduced to the set of
414 images containing corrosion. The resultant dataset, comprising 49 images,
415 was evaluated using the five defect detectors, as well as for the WCCD algo-
416 rithm. Figure 9 shows the ROC and PR curves for the different detectors.
417 As can be observed, the ROC curve for WCCD is considerable below the
418 curves of all our defect detectors. Regarding the PR curves, WCCD attains
419 higher precision than the contrast and symmetry-based detectors for certain
420 values of recall, although the WCCD curve is below the curves corresponding
421 to the three combined detectors (i.e OR, ORA and AND).

422 In a second comparative assessment, we have used the defect detector
423 presented in Bonnin-Pascual and Ortiz (2014a). This algorithm combines
424 contrast and symmetry information through the Bayesian framework SUN
425 (Zhang et al., 2008) to provide a saliency value for every pixel in the image.
426 To be more precise, the saliency at a given pixel z is defined as:

$$S_z = \frac{1}{p(F = f_z)} p(F = f_z | C = 1), \quad (18)$$

427 where F represents the visual features associated to a pixel (contrast and/or
428 symmetry in our case), f_z represents the feature values observed at z , and
429 C denotes whether a pixel belongs to the target class or not ($1 =$ defective

430 area). Using this formulation, the saliency of a given pixel z decreases as the
431 probability of feature f_z gets higher, and increases as the probability of f_z
432 in defects increases. The Parzen windows method was applied once again to
433 estimate those probabilities, using all the images of the dataset.

434 Notice that, despite both approaches use contrast and symmetry as fea-
435 tures to describe the defective areas, the SUN-based detector requires from
436 a sort of training stage to estimate the probability distributions, and its fea-
437 ture combination is performed within a probabilistic formulation while, in
438 this work, we propose three different combinations inspired by logical oper-
439 ators.

440 To perform the assessment, we have used the complete dataset. Three
441 different configurations of the SUN-based detector have been considered: us-
442 ing only contrast, using only symmetry and using both features. These three
443 configuration have been evaluated through Leave-One-Out-Cross-Validation
444 (Duda et al., 2000) and their corresponding ROC/PR curves and AUC val-
445 ues have been computed. Figure 10 compares these results with the ones
446 obtained for the corresponding three configurations of our framework: us-
447 ing only contrast, using only symmetry and using both features combined
448 through the OR operator.

449 As can be observed, the results obtained for the defect detection frame-
450 work presented in this paper are very similar to the ones obtained using the
451 SUN framework. This indicates that a successful defect detection can be
452 attained using contrast and symmetry information without performing any
453 training stage, as the SUN-based detectors do.

454 In a fourth kind of experiment, we have checked the usability of the

455 defect detectors with images taken by the aerial robotic platform presented
456 in Bonnin-Pascual et al. (2015). Regarding vessel inspection, this platform
457 can be operated under *inspection mode*, what means, during image capture,
458 constant and reduced speed (if it is not hovering) while keeping the same
459 distance and orientation with regard to the front wall, to improve image
460 quality.

461 The aerial vehicle was flown in three different compartments of a bulk-
462 carrier: a cargo hold, a top-side ballast tank, and the fore-peak tank. The
463 operating conditions in each compartment were very different. On the one
464 hand, the metallic plates inside the cargo hold did not exhibit corrosion which
465 could be visually appreciated. On the other hand, the metallic surfaces inside
466 the top-side and fore-peak tanks did present several corroded areas. Three
467 datasets have been generated (one for each compartment) containing a total
468 amount of 220 images. Ground-truth images have been manually produced
469 in the same way as for previous datasets.

470 To perform the assessment, the performance of our five defect detectors
471 has been evaluated considering the datasets of the vessel compartments af-
472 fected by corrosion. To be precise, Fig. 11 provides the results obtained for
473 the top-side tank dataset, while Fig. 12 shows the curves corresponding to the
474 fore-peak tank dataset. Additionally, Fig. 13 provides the metrics resulting
475 when considering both ballast tanks.

476 Unlike what happened with the original dataset, symmetry outperforms
477 contrast for the datasets comprising images taken by the MAV in the top-side
478 or/and the fore-peak tanks. On the one hand, the ROC curve obtained for
479 symmetry attains a position closer to the (0, 1) corner than the curve ob-

480 tained for contrast. On the other hand, symmetry results into considerably
481 higher values of precision. Regarding the combined methods, the OR and
482 AND combinations yield very similar results, and outperform all the other
483 single and combined versions of the detector. Nevertheless, the ORA combi-
484 nation presents slightly poorer performance due to the excessive importance
485 given to the different channels of contrast: intensity, colour and orientation.

486 A final experiment has been performed including also the images from the
487 cargo hold dataset, so that all the images from the three vessel compartments
488 have been considered. The performance metrics for this experiment can be
489 found in Fig. 14. Remember that the images from the cargo hold do not
490 present corroded areas, so that any positive detection dramatically increases
491 the FPR and decreases the precision. In the ROC space, the three combined
492 methods again provide better performance than the single-feature detectors.
493 Regarding the PR curve, the OR and AND combinations outperform all the
494 other versions of the detector, while the precision of the ORA combination is
495 more reduced due to the poorer performance provided by the contrast-based
496 method.

497 By way of example, Fig. 15 shows an image for each dataset, together
498 with its ground truth and the outputs provided by the five defect detectors.

499 In a last experiment, the vehicle was flown in front of a 2.5×4 m surface
500 containing corroded areas while the vision system was taking pictures at 10
501 Hz. The collected images were then processed by the image mosaicing algo-
502 rithm described in Garcia-Fidalgo et al. (2015), which managed to produce
503 the seamless composite shown in Fig. 16 (A). Finally, the mosaic was anal-
504 ysed using the five defect detectors, which provided the defect maps shown in

505 Fig. 16 (C-G). Notice that the detector does not analyse the mosaic borders
506 since contrast and symmetry levels can not be successfully computed in these
507 areas. A ground truth image has been manually generated for the mosaic
508 (see Fig. 16 (B)) in order to check the quality of the defect maps. As can be
509 observed, lighter pixels in the defect maps, that is, those which are likelier
510 to correspond to defects according to our detectors, are indeed labelled in
511 white in the ground truth.

512 **5. Conclusions**

513 A novel approach for defect detection on vessel structures has been pre-
514 sented. This has been devised as a generic framework that can be configured
515 ad hoc, selecting the features (and the way to combine them) that pro-
516 vide a more successful classification of the defective and non-defective areas.
517 The detection framework can merge multi-scale information of the selected
518 features to increase the robustness of the detection against changes in the
519 distance to the inspected area while collecting the images.

520 The selection of the features for our particular problem has been inspired
521 by the idea of conspicuity and taking into account the kind of defects that
522 appear in the metallic structures of vessels. Contrast in intensity, color and
523 orientation, and isotropic symmetry have been the features selected. Three
524 different combinations of these features inspired by logical operators have
525 also been considered, in order to merge the information they convey and
526 provide a better description of the defective situations.

527 The different defect detectors have shown good classification performances,
528 improving the results obtained from previous detectors. In comparison with

529 them, the presented approach does not require from tuning a large set of
530 working parameters nor performing a previous training stage.

531 Regarding the feature set, the results obtained for the different datasets
532 indicate that contrast and symmetry complement each other, so that one can
533 provide the proper information to discriminate whether an area is defective
534 or not when the other feature maybe fails, and vice versa.

535 The usability of the proposed solution has also been proved using images
536 collected by a micro-aerial robotic platform devised for vessel inspection,
537 which has been flown in different areas inside a bulk carrier.

538 The experimental results have shown that the algorithm is also able to
539 successfully detect defective areas in mosaics generated from these images.
540 During a vessel inspection campaign, the use of mosaics allows us to extract
541 more information about the state of the inspected surface since defective
542 areas are not split over multiple images.

543 **Acknowledgements**

544 This work is partially supported by FEDER funding, by project nr.
545 AAEE50/2015 (Direcció General d’Innovació i Recerca, Govern de les Illes
546 Balears), and by project INCASS. The latter has received research funding
547 from the EU FP7 under GA 605200. This publication reflects only the au-
548 thors views and the European Union is not liable for any use that may be
549 made of the information contained therein.

550 **References**

- 551 Avraham, T., Lindenbaum, M., 2010. Esaliency (Extended Saliency): Mean-
552 ingful Attention Using Stochastic Image Modeling. *IEEE Transactions on*
553 *Pattern Analysis and Machine Intelligence* 32 (4), 693–708.
- 554 Bonnin-Pascual, F., 2010. Detection of Cracks and Corrosion for Automated
555 Vessels Visual Inspection. Master’s thesis, University of the Balearic
556 Islands.
557 URL {http://dmi.uib.es/~xbonnin/static/papers/mthesis_
558 [Bonnin2010.pdf](http://dmi.uib.es/~xbonnin/static/papers/mthesis_Bonnin2010.pdf)}
- 559 Bonnin-Pascual, F., Ortiz, A., 2014a. A Probabilistic Approach for Defect
560 Detection Based on Saliency Mechanisms. In: *IEEE International Confer-*
561 *ence on Emerging Technologies and Factory Automation*.
- 562 Bonnin-Pascual, F., Ortiz, A., 2014b. Corrosion Detection for Automated Vi-
563 sual Inspection. In: Aliofkhazraei, D. M. (Ed.), *Developments in Corrosion*
564 *Protection*. InTech, Ch. 25, pp. 619–632.
- 565 Bonnin-Pascual, F., Ortiz, A., Garcia-Fidalgo, E., Company, J. P., 2015. A
566 Micro-Aerial Platform for Vessel Visual Inspection based on Supervised
567 Autonomy. In: *IEEE/RSJ International Conference on Intelligent Robots*
568 *and Systems*. pp. 46–52.
- 569 Borji, A., Ahmadabadi, M. N., Araabi, B. N., 2011. Cost-sensitive Learn-
570 ing of Top-Down Modulation for Attentional Control. *Machine Vision and*
571 *Applications* 22 (1), 61–76.

- 572 Borji, A., Itti, L., 2013. State-of-the-Art in Visual Attention Modeling. *IEEE*
573 *Transactions on Pattern Analysis and Machine Intelligence* 35 (1), 185–207.
- 574 Chin, R. T., Harlow, C. A., 1982. Automated Visual Inspection: A Survey.
575 *IEEE Transactions on Pattern Analysis and Machine Intelligence* 4 (6),
576 557–573.
- 577 Duda, R. O., Hart, P. E., Stork, D. G., 2000. *Pattern Classification*, 2nd
578 Edition. Wiley Interscience.
- 579 Eich, M., Bonnin-Pascual, F., Garcia-Fidalgo, E., Ortiz, A., Bruzzone, G.,
580 Koveos, Y., Kirchner, F., 2014. A Robot Application to Marine Vessel
581 Inspection. *Journal of Field Robotics* 31 (2), 319–341.
- 582 Fawcett, T., 2006. An Introduction to ROC Analysis. *Pattern Recognition*
583 *Letters* 27 (8), 861–874.
- 584 Garcia-Fidalgo, E., Ortiz, A., Bonnin-Pascual, F., Company, J. P., 2015.
585 A Mosaicing Approach for Vessel Visual Inspection using a Micro-Aerial
586 Vehicle. In: *IEEE/RSJ International Conference on Intelligent Robots and*
587 *Systems*.
- 588 Itti, L., Koch, C., Niebur, E., nov 1998. A Model of Saliency-based Visual At-
589 tention for Rapid Scene Analysis. *IEEE Transactions on Pattern Analysis*
590 *and Machine Intelligence* 20 (11), 1254 –1259.
- 591 Jahanshahi, M. R., Kelly, J. S., Masri, S. F., Sukhatme, G. S., 2009. A
592 Survey and Evaluation of Promising Approaches for Automatic Image-
593 based Defect Detection of Bridge Structures. *Structure and Infrastructure*
594 *Engineering* 5 (6), 455–486.

- 595 Koch, C., Ullman, S., 1985. Shifts in Selective Visual Attention: Towards
596 the Underlying Neural Circuitry. *Human Neurobiology* 4 (4), 219–227.
- 597 Kootstra, G., Nederveen, A., Boer, B. D., 2008. Paying Attention to Sym-
598 metry. In: *British Machine Vision Conference*. BMVA Press, pp. 111.1–
599 111.10.
- 600 Kootstra, G., Schomaker, L., 2009. Prediction of Human Eye Fixations Using
601 Symmetry. In: *Annual Conference of the Cognitive Science Society*. pp.
602 56–61.
- 603 Li, J., Tian, Y., Huang, T., Gao, W., 2010. Probabilistic Multi-task Learning
604 for Visual Saliency Estimation in Video. *International Journal of Computer*
605 *Vision* 90 (2), 150–165.
- 606 Malamas, E. N., Petrakis, E. G. M., Zervakis, M., Petit, L., Legat, J.-D.,
607 2003. A Survey on Industrial Vision Systems, Applications and Tools. *Im-*
608 *age and Vision Computing* 21, 171–188.
- 609 Mumtaz, M., Masoor, A. B., Masood, H., 2010. A new approach to aircraft
610 surface inspection based on directional energies of texture. In: *Interna-*
611 *tional Conference on Pattern Recognition*. pp. 4404–4407.
- 612 Newman, T. S., 1995. A Survey of Automated Visual Inspection. *Computer*
613 *Vision and Image Understanding* 61 (2), 321–262.
- 614 Oullette, R., Browne, M., Hirasawa, K., 2004. Genetic Algorithm Optimiza-
615 tion of a Convolutional Neural Network for Autonomous Crack Detection.
616 In: *Congress on Evolutionary Computation*. pp. 516–521.

- 617 Ozog, P., Eustice, R. M., 2015. Identifying Structural Anomalies in Image
618 Reconstructions of Underwater Ship Hulls. In: IEEE/MTS OCEANS Con-
619 ference and Exhibition.
- 620 Petricca, L., Moss, T., Figueroa, G., Broen, S., 2016. Corrosion Detection
621 using A.I.: A Comparison of Standard Computer Vision Techniques and
622 Deep Learning Model. In: International Conference on Computer Science,
623 Engineering and Information Technology. pp. 91–99.
- 624 Siegel, M., Gunatilake, P., 1998. Remote enhanced visual inspection of air-
625 craft by a mobile robot. In: IEEE International Workshop on Emergent
626 Technologies, Intelligent Measurement and Virtual.
- 627 Theodoridis, S., Koutroumbas, K., 2006. Pattern Recognition, 3rd Edition.
628 Academic Press.
- 629 Wolfe, J. M., 2007. Integrated Models of Cognitive Systems. In: Gray, W. D.
630 (Ed.), Guided search 4.0. Oxford University Press New York, NY, Ch. 8,
631 pp. 99–119.
- 632 Xie, X., 2008. A Review of Recent Advances in Surface Defect Detection
633 Using Texture Analysis Techniques. *Electronic Letters on Computer Vision
634 and Image Analysis* 7 (3), 1–22.
- 635 Yamaguchi, T., Hashimoto, S., 2010. Fast Crack Detection Method for Large-
636 size Concrete Surface Images Using Percolation-based Image Processing.
637 *Machine Vision and Applications* 21 (5), 797–809.
- 638 Zhang, L., Qiu, B., Yu, X., Xu, J., 2015. Multi-scale Hybrid Saliency Analysis

- 639 for Region of Interest Detection in Very High Resolution Remote Sensing
640 Images. *Image and Vision Computing* 35, 1–13.
- 641 Zhang, L., Tong, M. H., Marks, T. K., Shan, H., Cottrell, G. W., 2008. SUN:
642 A Bayesian Framework for Saliency Using Natural Statistics. *Journal of*
643 *Vision* 8 (7), 1–20.
- 644 Zhang, L., Yang, F., Zhang, Y. D., Zhu, Y. J., 2016. Road Crack Detec-
645 tion using Deep Convolutional Neural Network. In: *IEEE International*
646 *Conference on Image Processing*. pp. 3708–3712.

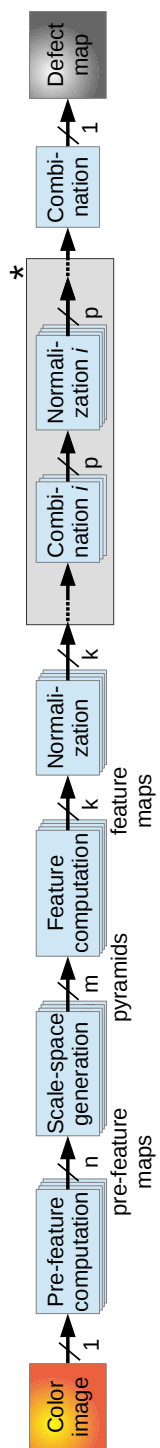


Figure 1: Generic framework for defect detection. (\star) means zero or more than zero instances of the corresponding stage.

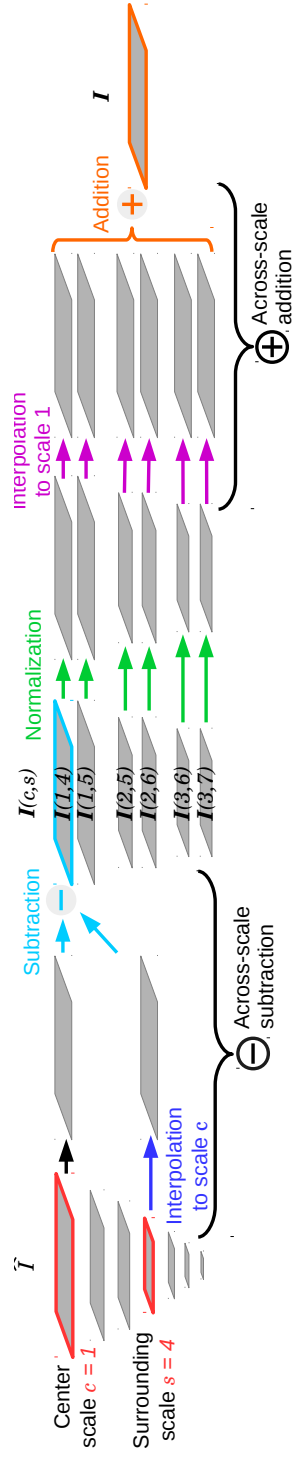


Figure 2: Illustration of feature map computation: case of intensity-contrast map.

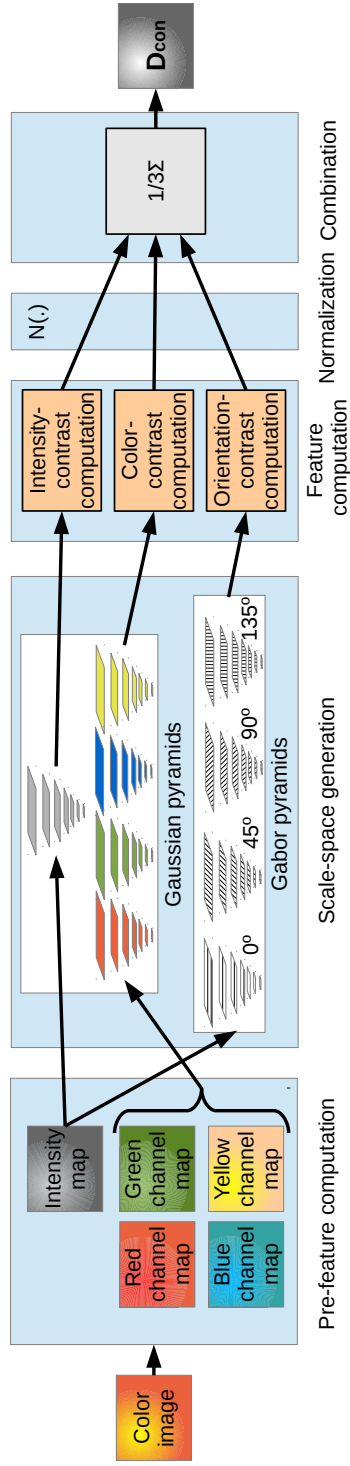


Figure 3: Implementation of the contrast-based defect detector using the generic framework.

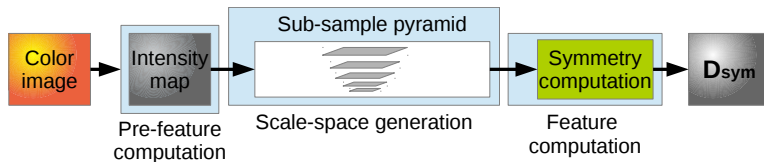


Figure 4: Implementation of the symmetry-based defect detector using the generic framework.

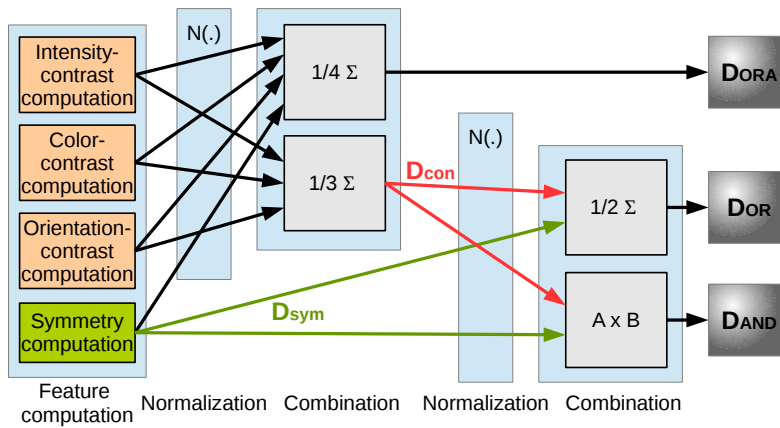


Figure 5: Set up of the normalization and combination stages for the defect detectors merging contrast and symmetry information.

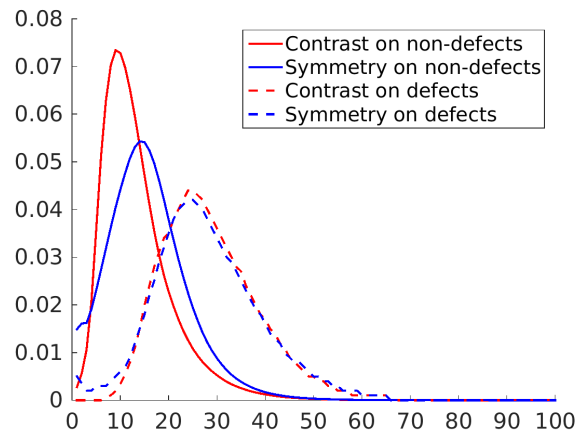


Figure 6: Estimated PDFs for contrast and symmetry features.

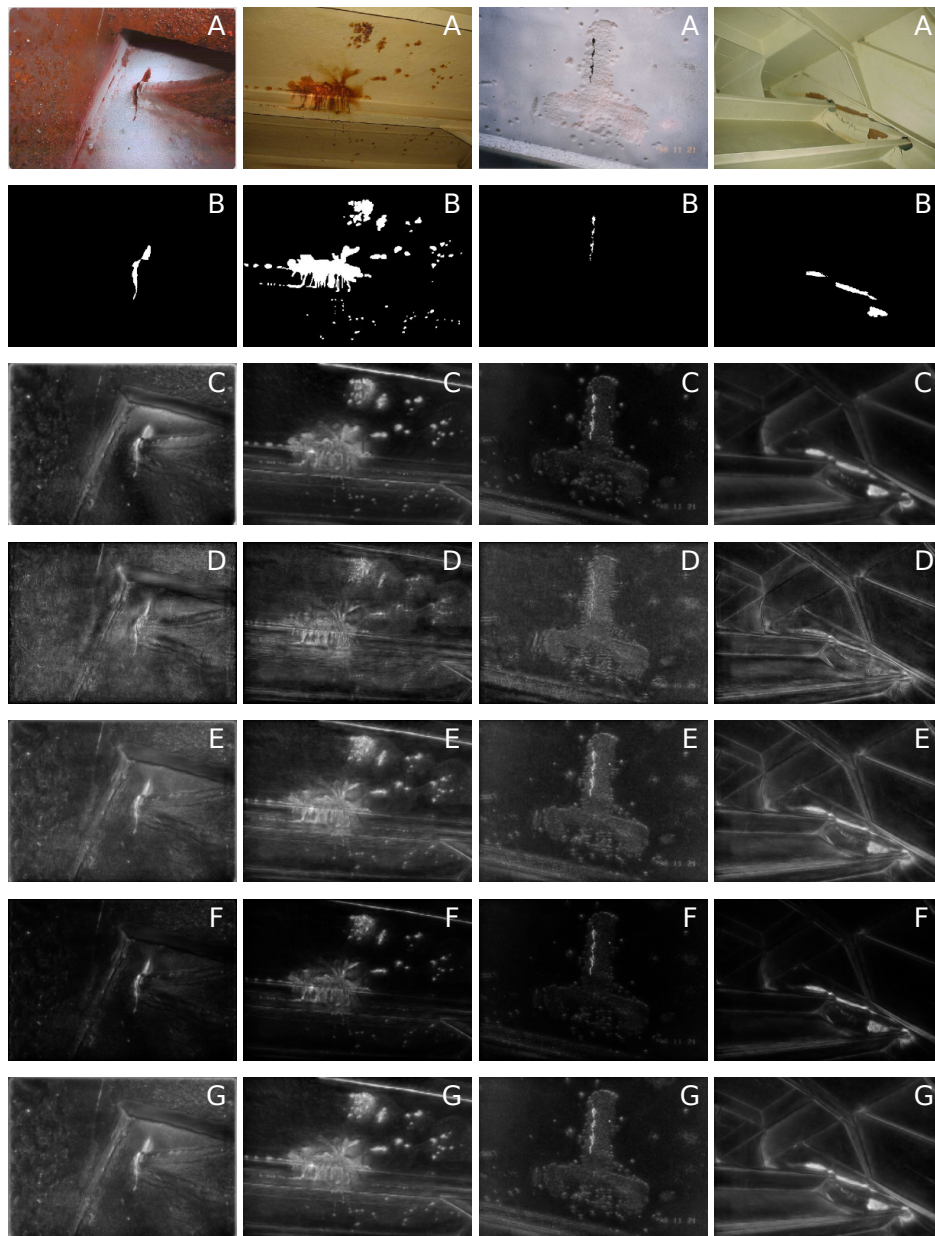


Figure 7: Test images with their associated ground truth and resulting defect maps. A: original image. B: ground truth. C and D: respectively, defect maps obtained using contrast and symmetry. E, F and G: respectively, defect maps obtained from the OR, AND and ORA combinations.

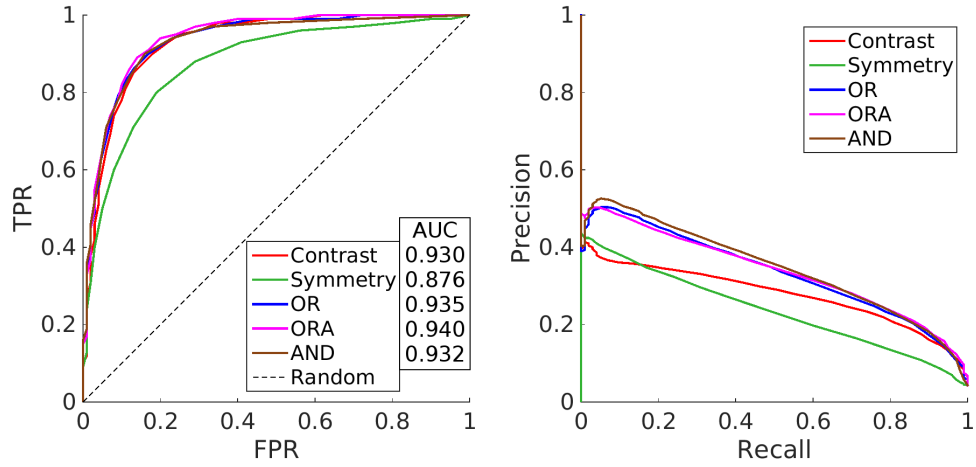


Figure 8: Performance of the five defect detectors: (left) ROC curves and AUC values, (right) PR curves.

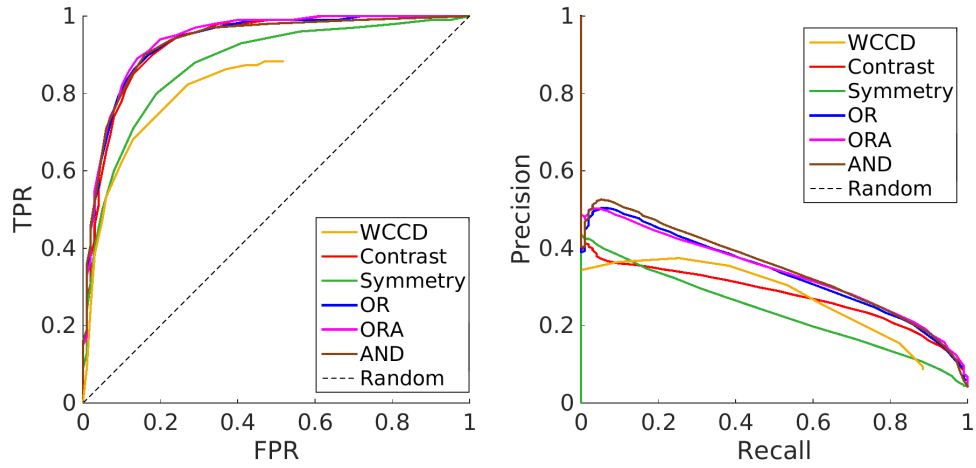


Figure 9: Comparison between our defect detectors and the WCCD algorithm: (left) ROC curves and AUC values, (right) PR curves.

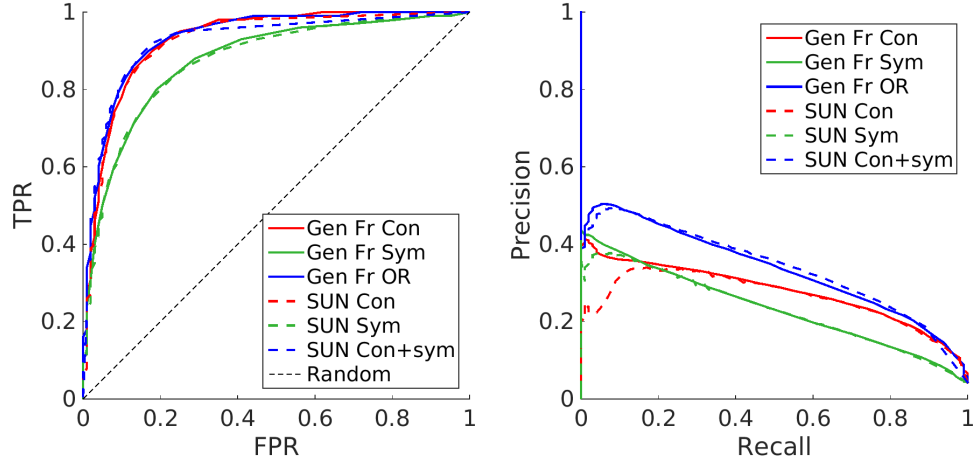


Figure 10: Comparison among the defect detectors presented in this work (i.e. using the generic framework) and the SUN-based detectors: (left) ROC curves and AUC values, (right) PR curves.

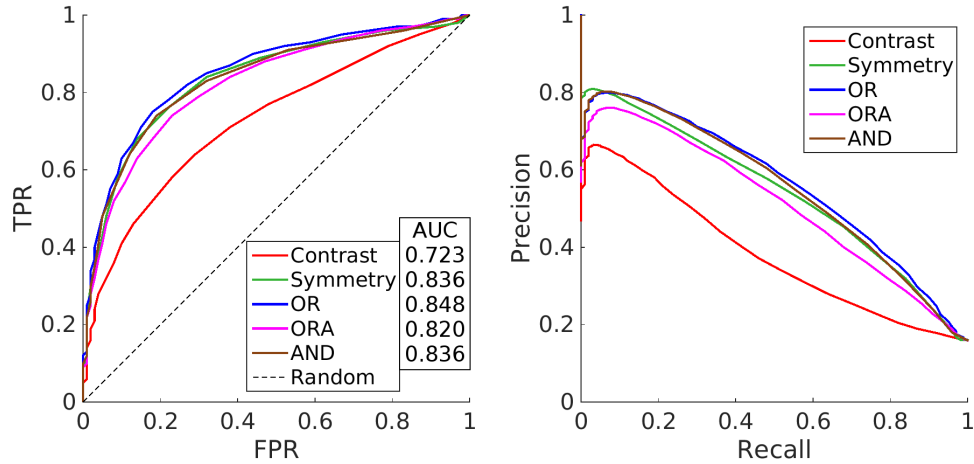


Figure 11: Performance of the five detectors evaluating images taken from a top-side ballast tank inside a bulk carrier: (left) ROC curves and AUC values, (right) PR curves.

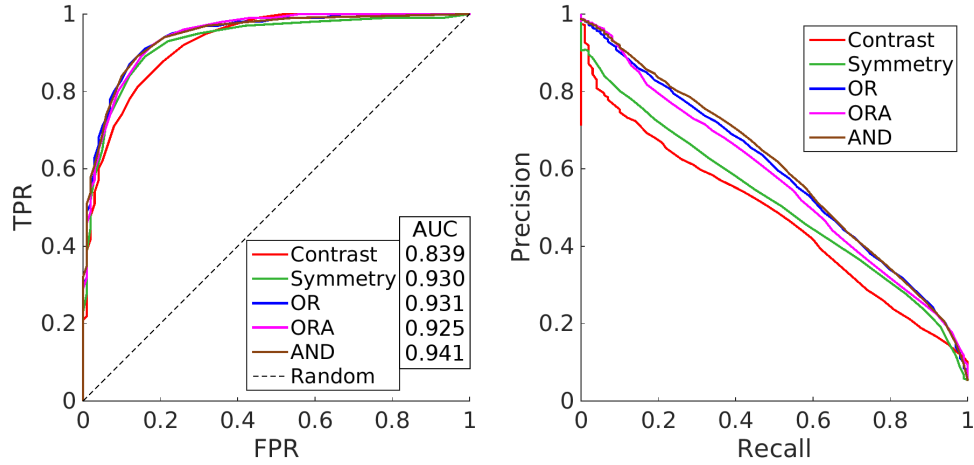


Figure 12: Performance of the five detectors evaluating images taken from the fore-peak tank of a bulk carrier: (left) ROC curves and AUC values, (right) PR curves.

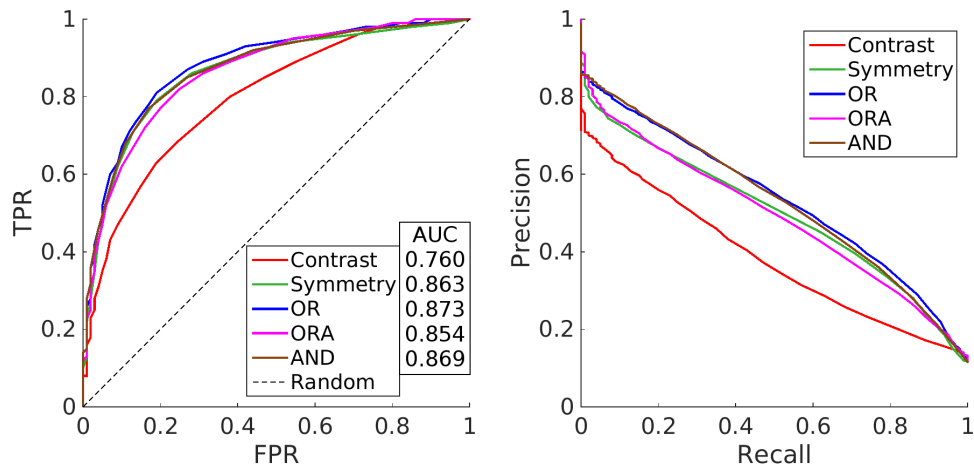


Figure 13: Performance of the five detectors evaluating images taken from a top-side and the fore-peak tanks of a bulk carrier: (left) ROC curves and AUC values, (right) PR curves.

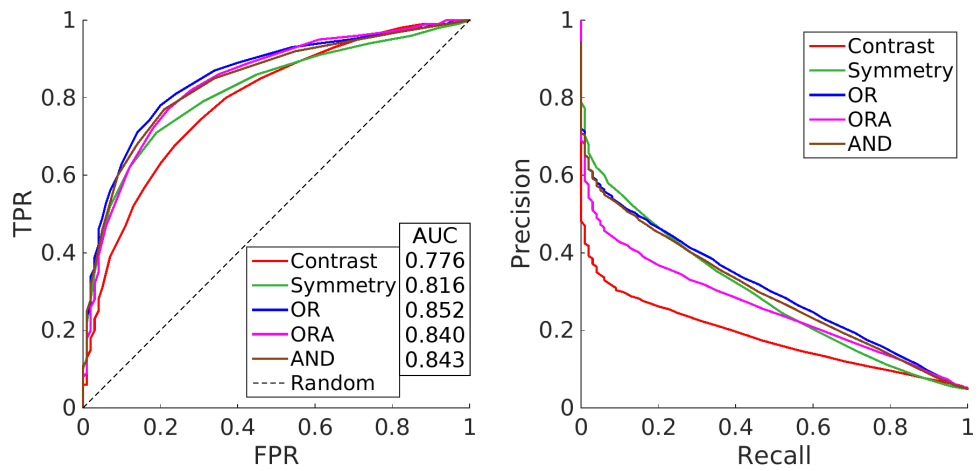


Figure 14: Performance of the five detectors evaluating the images taken from three different spaces inside a bulk carrier, namely a cargo hold, a top-side ballast tank and the fore-peak tank: (left) ROC curves and AUC values, (right) PR curves.

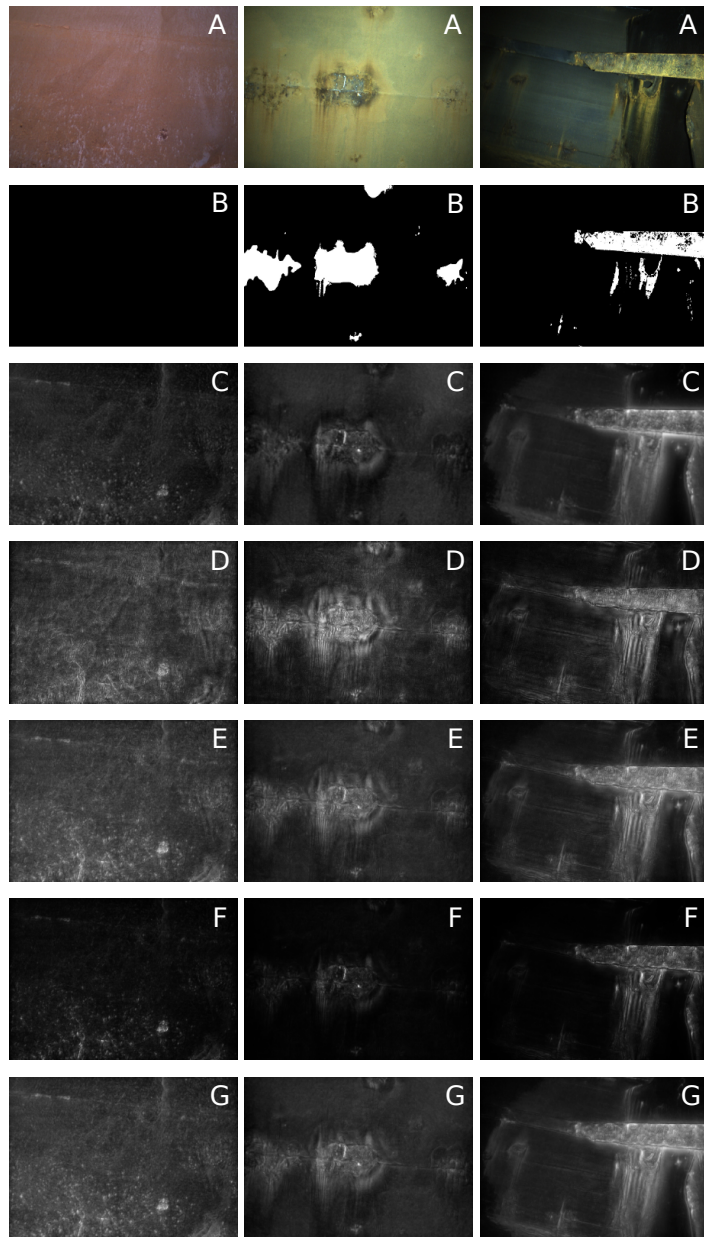


Figure 15: Results for three images taken inside a bulk carrier: a cargo hold (left), a top-side ballast tank (middle), and the fore-peak tank (right). A: original image. B: ground truth. C and D: respectively, defect maps obtained using contrast and symmetry. E, F and G: respectively, defect maps obtained from the OR, AND and ORA combinations.

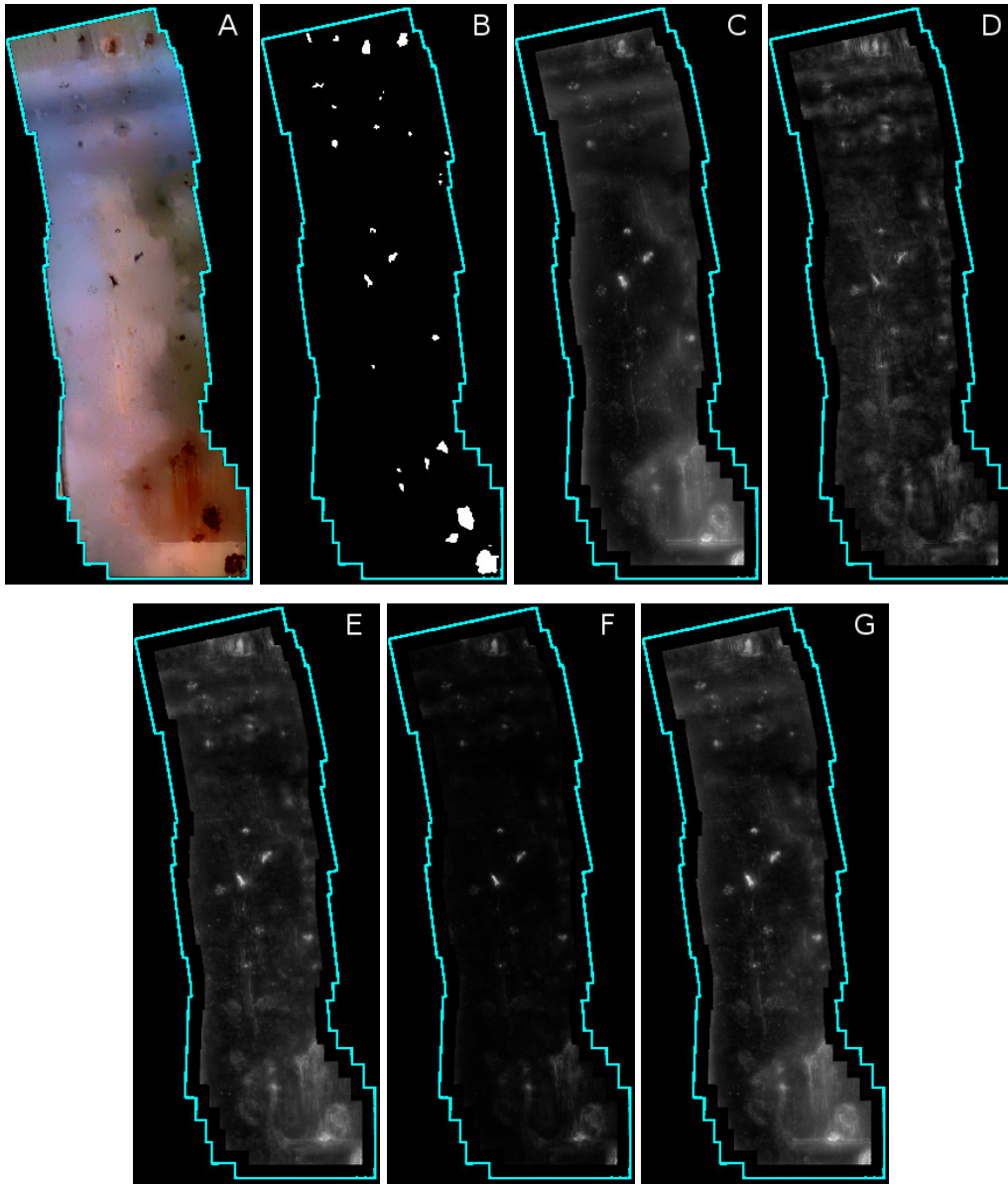


Figure 16: Detection results when inspecting an image mosaic. A: mosaic built from images collected by the aerial vehicle. B: ground truth image manually generated. C and D: respectively, defect maps obtained using contrast and symmetry. E, F and G: respectively, defect maps obtained from the OR, AND and ORA combinations. In defect maps, lighter pixels are likelier to correspond to defects (mosaic borders are not analysed).