



HAL
open science

One-class SVM based outlier detection strategy to detect thin interlayer debondings within pavement structures using Ground Penetrating Radar data

Shreedhar Savant Todkar, Vincent Baltazart, Amine Ihamouten, Xavier Dérobert, David Guilbert

► To cite this version:

Shreedhar Savant Todkar, Vincent Baltazart, Amine Ihamouten, Xavier Dérobert, David Guilbert. One-class SVM based outlier detection strategy to detect thin interlayer debondings within pavement structures using Ground Penetrating Radar data. *Journal of Applied Geophysics*, 2021, 192, pp.104392. 10.1016/j.jappgeo.2021.104392 . hal-04092120

HAL Id: hal-04092120

<https://hal.science/hal-04092120v1>

Submitted on 22 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

One-class SVM based outlier detection strategy to detect thin interlayer debondings within pavement structures using Ground Penetrating Radar data

Shreedhar Savant Todkar^{a,*}, Vincent Baltazart^a, Amine Ihamouten^b, Xavier Dérobert^c,
Cédric Le Bastard

^aDepartment of Components and Systems (COSYS - SII), University Gustave Eiffel, Nantes Campus, 44344 Bouguenais, France.

^bDepartment of Materials and Structures (MAST - LAMES), University Gustave Eiffel, Nantes Campus, 44344 Bouguenais, France.

^cAssessment and Imaging Laboratory (GERS - GeoEND), University Gustave Eiffel, Nantes Campus, 44344 Bouguenais, France.

Abstract

In this paper, we present a processing method to detect millimeter interlayer debondings from Ground Penetrating Radar (GPR) B-scan images. The method is matched to carry out rapid debonding detection at the operational level. A machine learning based outlier-detection strategy namely, One-class Support Vector Machines (OCSVM) is proposed to detect A-scan data vectors which differ from a reference data set collected over a known healthy pavement area.

OCSVM is tested on both simulated and experimental data representing GPR data over various artificial millimetric debondings at 2.6 GHz and 4.2 GHz from respectively ground-coupled and air-coupled radar configurations. The experimental data were collected at the Accelerated Pavement Test site located in the Nantes campus of Université Gustave Eiffel. The simulated models on the other hand were generated using a numerical EM solver based on Finite Difference Time Domain (FDTD) method namely, GprMax. Simulation tests allow to conduct sensitivity analysis to determine the robustness of the detection method at various signal-to-noise ratios (10 dB to 60 dB).

The proposed OCSVM method demonstrated high performance on both simulated and experimental data to detect thin interlayer debondings over various GPR configurations.

Keywords: Ground Penetrating Radar (GPR), thin debondings, One-class SVM (OCSVM), Air-coupled radar, Ground-coupled radar, Sensitivity Analysis, Finite Difference Time Domain (FDTD), GprMax.

1. Introduction

In the context of pavement monitoring, an early stage defect detection is of great importance to maintain its durability, conduct Life Cycle Assessments (LCA) and avoid significant failures (Büyüköztürk and Yu, 2006; Hajj et al., 2010). However, with the

*Corresponding author

Email addresses: shreedhar.todkar@univ-eiffel.fr (Shreedhar Savant Todkar),
vincent.baltazart@univ-eiffel.fr (Vincent Baltazart),
amine.ihamouten@univ-eiffel.fr (Amine Ihamouten),
xavier.derobert@univ-eiffel.fr (Xavier Dérobert), cedric.lebastard@gmail.com
(Cédric Le Bastard)

37 passage of time and heavy traffic, extreme variations in weather conditions and water
38 seepage, two types of pavement damages tend to occur: Surface and subsurface defects
39 (Wright et al., 2014). In this paper, we focus on the latter type of defects, especially
40 horizontal inter-layer cracks. The inter-layer cracks, called debondings (or delamina-
41 tions), usually appear between the top two layers of the pavement (Todkar et al., 2017).
42 These defects may give rise to stripping and/or reflexive cracks that lead to open sur-
43 face cracks severely degrading the structural integrity of the pavement. Non-destructive
44 testing (NDT) is one such technique that has been widely used in this aspect.
45 Ground Penetrating Radar (GPR) is one well-known NDT electromagnetic method
46 used in the field of Civil engineering for Non-destructive Testing (NDT) applications.
47 (Wai-Lok Lai et al., 2018) provides an overview of the 30-year journey of GPR tech-
48 nology in civil engineering. Owing to its non-invasive capability and the sensitivity
49 to material permittivity, GPR has been used in the literature for various applications
50 including water content measurement (Kaplanvural et al., 2018), rebar corrosion in
51 concrete (Zaki et al., 2018), pavement layer thickness estimation (Pinel et al., 2009),
52 subsurface defects (Dong et al., 2016), anti-personnel mine detection (Ebrahim et al.,
53 2018; Tellez and Scheers, 2017) *etc.*

54 In this paper, we focus on the detection of the horizontally stratified thin debondings
55 that occur at the interface between the top two layers of the pavement structures. How-
56 ever as mentioned in (Benedetto et al., 2017), GPR data require suitable processing
57 techniques that can interpret and help in the decision making process (such as clas-
58 sification, identification *etc.*). This has motivated the development of specific signal
59 processing techniques to conduct early detection of thin debondings.

60 Support Vector Machines (SVM) are generally implemented as they pose low compu-
61 tational complexity, since they are based on support vectors (Tbarki et al., 2017b). In
62 addition, the convexity of the optimal problem characterizes the unique solution ob-
63 tained using SVM. Two-class SVM has been used in (Todkar et al., 2019, 2018, 2017)
64 on GPR data to classify data vectors into defective (debonding) and non-debonding
65 (healthy) A-scan classes.

66 Although the Two-class classification provided some impressive results in these cases,
67 it lacked some key aspects. Firstly, it required *a priori* knowledge about the two data
68 classes during the learning step, which, is not always available (Salem and Stolfo,
69 2012). Secondly, the imbalance in the two data (debonding and non-debonding) classes
70 during the learning step results in a lower performance of the method. Two-class SVM
71 also generally demands a larger learning dataset for best performance as shown in (Tod-
72 kar et al., 2019).

73 An alternative strategy is proposed here to detect debondings within pavement struc-
74 tures using One-class SVM (OCSVM). A unique reference data class is used to identify
75 outlier subsurface pavement conditions. In this context, the outlier is defined as the A-
76 scan vectors that deviate from the reference data set collected over healthy pavement
77 sections. The proposed strategy is expected to be better matched to both the short-term
78 condition of field test and the limited prior knowledge which is available at the opera-
79 tional level. In the literature, OCSVM method has been used along with GPR to detect
80 buried landmines (Tbarki et al., 2017b, 2018). Within this scope, the machine learning
81 OCSVM (Schölkopf et al., 2001) is applied for detecting outlier radar data vectors (*i.e.*
82 A-scans). OCSVM is advantageous as it performs faster training and testing (Salem

83 and Stolfo, 2012) due to its simplistic classification model. It also eliminates the need
84 for labeled two-class information during learning stage (Senf et al., 2006). Besides, as
85 mentioned in (Tbarki et al., 2017a), the outlier data (*i.e.*, debonding data in our appli-
86 cation) are either not readily available or they very costly to acquire.
87 The versatility of the proposed approach is illustrated with the data created using the
88 Finite-Difference Time Domain (FDTD) method simulating two radar configurations:
89 air-coupled and ground-coupled GPR data collected over the same pavement structure
90 embedded with artificial debondings. OCSVM is thereafter formulated using some lo-
91 cal signal features to provide the required refined signal analysis of the time domain
92 GPR signatures. The method demonstrated a quick and efficient defect detection and
93 classification for both simulated and experimental data over ground-coupled and air-
94 coupled radar configurations.
95 In this paper, we present the experimental and simulated GprMax datasets used to test
96 the proposed detection OCSVM method, and describe its approach. The simulated
97 data are modeled using Ground-Coupled-GPR (GC-GPR) and Air-coupled Stepped
98 Frequency GPR (SF-GPR) configurations. Throughout, OCSVM parameters are ex-
99 tensively tested by Sensitivity Analysis using GprMax data, followed by demonstrating
100 its performance on experimental data.

101 **2. Datasets for method tests and validation**

102 In this paper, two types of datasets are used to test the proposed detection method: a
103 pavement model generated using GprMax, and experimental data collected at our uni-
104 versity’s Accelerated Pavement Test (APT) site. Both datasets are acquired/generated
105 using the two configurations: ground-coupled and air-launched GPR. We now present
106 these two data sets.

107 *2.1. Experimental datasets*

108 To validate the OCSVM method over experimental data, the data is collected at
109 IFSTTAR’s* APT site (Simonin et al., 2012) (currently known as University Gustave
110 Eiffel).

111 *2.1.1. The Fatigue Carousel*

112 The fatigue carousel situated at the Nantes campus of University Gustave Eiffel is
113 a 120 m long and 6 m wide one off-circular outdoor APT test track. It facilitates sev-
114 eral loading stages (65 kN on single wheel, twin wheel, tandem or tridem axles) with
115 various configurations and rotational speeds of up to 100 km h⁻¹ which reproduces the
116 effect of in-use traffic at an accelerated rate.

117 Of its 120 m test track, the site dedicates a 25 m section for our experiments. Here, a
118 two-layered structure with the top layer (course layer) of about 5 cm to 6 cm thickness
119 and the base layer of about 8 cm thickness over a granular sub-base are laid.

120 Three artificial rectangular defects of various thickness and material compositions are

*French Institute of Science and Technology for Transport, Development and Networks

121 embedded in this section (ref. Figure 1a). These defects namely, Geotextile, Sand and
 122 Tack-free based defects are monitored over several loading stages of 10K, 50K and
 123 100K (where K indicates one thousand loading cycles/rotations of the fatigue carousel
 124 (Dérobert and Baltazart, 2012)). Geotextile and Sand based defects are defects with
 125 larger thicknesses presenting the highly degraded pavement conditions. Tack-free de-
 126 fects represent the lack of coating conditions between the pavement layers and as such,
 127 they closely resemble a realistic defective pavement structure with thin debondings.
 128 Thus, in this paper we present the debonding detection for the weakest defect type,
 129 namely, Tack-free. A coring test at the end of the pavement life-cycle showed that the
 130 maximum debonding thickness of Tack-free defect zone was within the range of 2 mm
 131 to 3 mm. Figure 1a shows the embedded defects before laying the course layer. The
 132 zone I13 represents the Tack-free based defects (Simonin et al., 2016, 2012). For our
 133 experiments, this zone is used for data acquisition. The areas marked ‘Zone I11’ and
 134 ‘Zone I12’ respectively represent Geotextile and Sand-based defects which are not in-
 135 cluded in this paper. (Dérobert et al., 2020) provides the recorded data for both radar
 136 types and the detailed description of the test track with the three embedded defects.
 137 The GPR probing pavement section consists of three layers (see Figure 1c): Layers 1
 138 and 2 are asphalt surface pavements (with $\varepsilon_{r1} < \varepsilon_{r2}$) and Layer 3 represents the arti-
 139 ficial debonding layer whose thickness ranges from 2 mm to 3 mm with $\varepsilon_{r1} < \varepsilon_{r2} <$
 140 $\varepsilon_{r,deb}$ for Tack-free defects. It should be noted that the permittivities may be reinforced
 141 by water seeping with time (Todkar et al., 2019).

142 2.1.2. Data collection

143 The data collection is done using the two main radar configurations which are used
 144 at the operational level. The first configuration is a ground-coupled commercial GPR
 145 namely, GSSI SIR-3000 (GSSI-GPR) operating at a peak magnitude centered at $f_c =$
 146 2.6 GHz in free space.

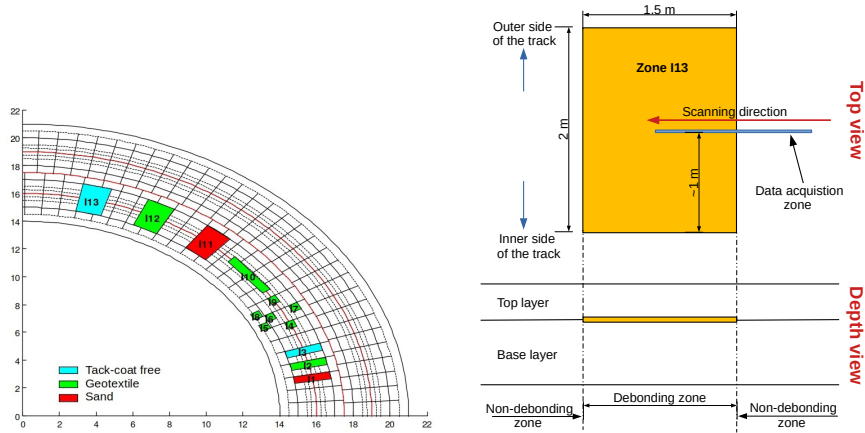
147 The GSSI-GPR uses a high-frequency quasi-monostatic transmitter-receiver (Tx - Rx)
 148 bow-tie antenna setup (Todkar et al., 2017). The radar allows data collection at traf-
 149 fic speed of up to 10 km h^{-1} . The radar is capable to acquire up to 8192 samples
 150 per A-scan. However during experiments, a preset value of 1024 samples per A-scan is
 151 chosen with a time window duration of 8 ns and the data is collected at average walking
 152 speed. The GSSI-GPR uses a very small Tx - Rx offset maintaining it in a quasi-mono-
 153 static configuration. Figure 2a shows the GSSI-GPR data acquisition setup at IFST-
 154 TAR’s APT site.

155 The second GPR is an experimental air-coupled Step-frequency radar (SFR) whose
 156 peak magnitude is centered at $f_c = 4.2 \text{ GHz}$ in free space with frequency bandwidth
 157 $BW = 0.8 \text{ GHz}$ to 10.8 GHz . The SFR uses an ultra-wideband exponentially tapered
 158 slot antennas (Diakité et al., 2015) and collects the data in frequency domain at limited
 159 speed thanks to a Vector Network Analyzer (VNA). The frequency data is converted
 160 to time domain data using inverse Fourier transform for processing by OCSVM. The
 161 SFR time range is limited to 6 ns and includes 5000 samples per A-scan. The SFR, has
 162 an offset of 0.2 m between the Tx and Rx and is thereby in a bi-static configuration.
 163 Figure 2b shows the SFR data acquisition setup.

164 Thanks to toolbox provided by (Nesbitt et al., 2019), the GSSI data is easily accessed
 165 for processing on the Python environment. On the other hand, an inverse Fourier trans-



(a) 25 m track with artificial defects before laying the wearing course layer. 'Zone 113' indicates Tack-free based defects (Zones 111 and 112 are respectively Geotextile and Sand-based defects)



(b) Location of each embedded defect sections. Zone '113' represents the Tack-free defects that was used for data acquisition (Dérobert et al., 2020)

Figure 1: APT Test track at Univ. Gustave Eiffel

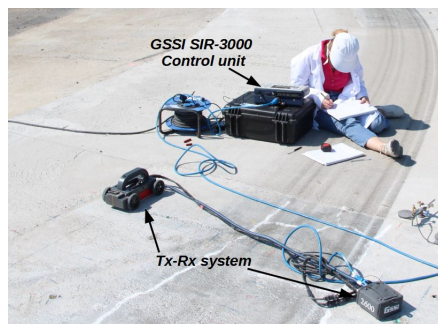
166 form is used on the SF-GPR data for processing.

167

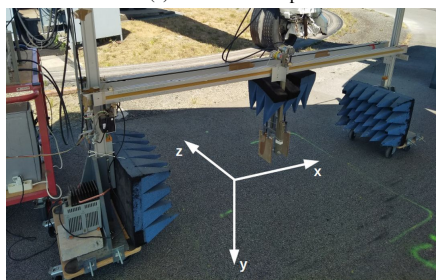
168 2.2. Numerical modeling using GprMax

169 For preliminary validation of the debonding detection method, we use an FDTD-
 170 based method that is used to generate subsurface models using EM wave propagation
 171 namely GprMax (Warren et al., 2015). Since we perform a 2D analysis, a 2D pavement
 172 model of dimensions $3.0 \text{ m} \times 0.28 \text{ m}$ is created. The domain size of the model is set
 173 to $3.0 \text{ m} \times 1.0 \text{ m}$. The spatial and temporal discretization sizes (or mesh size) used are
 174 respectively $\Delta x = 1 \times 10^{-3} \text{ m}$ and $\Delta y = 5 \times 10^{-4} \text{ m}$. The temporal resolution is
 175 set to $\Delta t = 1.45 \times 10^{-12} \text{ s}$ such that the stability condition known as CFL (Courant,
 176 Friedrichs, Lewy) condition is satisfied (Giannopoulos, 2005). The point source is a
 177 Hertzian dipole polarized in z -axis.

178 The pavement model used with GprMax is as shown in Figure 3. The model consists



(a) GSSI GPR setup



(b) Step-frequency GPR setup (Dérobert et al., 2020). The 'x'-axis (or spatial axis) represents the scanning direction, 'y'-axis (or time axis) is the depth and 'z'-axis is the lateral axis

Figure 2: Data acquisition setup at IFSTTAR's APT site

179 of two or three layers depending on the zone, which is divided into Non-debonding
 180 zone (two layers) and Debonding zone (three layers). The top course layer has relative
 181 permittivity of $\varepsilon_{r1} = 5$ and of thickness 0.08 m. The base layer on the other hand has
 182 permittivity of $\varepsilon_{r2} = 7$ with thickness 0.2 m.

183 In reality, an interlayer debonding may be composed of (a) air voids or water/moisture
 184 voids, (b) mixture of the pavement material from adjacent layers, or, (c) a combination
 185 of all of the above. In this paper, we focus on the detection of air-void defects. The
 186 debonding layer thus has a relative permittivity of $\varepsilon_{r,deb} = 2$ (value is used to incorpo-
 187 rate 'near air-void' defects) and the debonding thickness is assigned as $th_{deb} = 3$ mm
 188 (other values for th_{deb} have been also studied in the later sections).

189 In addition, it should be noted that whilst creating the model, the following assumptions
 190 were made:

- 191 1. Pavement layers are assumed to be homogeneous
- 192 2. Smooth surfaces and interfaces
- 193 3. Top layer is sufficiently thick to avoid overlapping of surface and interface echoes
- 194 4. Debonding thickness is smaller than the quarter of the wavelength (*i.e.*, $th_{deb} \leq$
 195 $\lambda_{deb}/4$, where λ_{deb} is the wavelength within the debonding layer)

196 The excitation waveform in the case of ground-coupled GPR is a Gaussian centered at
 197 $f_c = 2.6$ GHz whose antennas are positioned in contact with the surface (*i.e.*, $h_{ant} =$

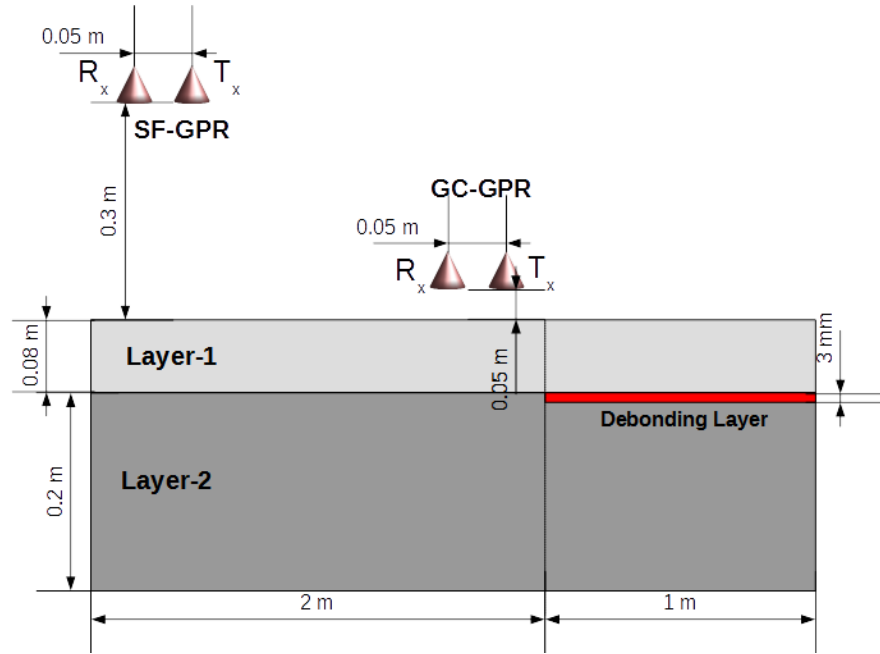


Figure 3: Generated pavement model using gprMax

198 0.05 m). These parameters are set to imitate the GSSI-GPR used in the experimental
 199 data (see Section 2.1). The distance between Tx and Rx is set to 0.05 m. For the
 200 air-launched GPR, the excitation waveform is the first Gaussian derivative centered at
 201 $f_c = 4.2$ GHz with $h_{ant} = 0.3$ m and the distance between Tx and Rx being 0.05 m.
 202 In both cases, the adjacent A-scans possess a spatial separation of 0.01 m. On the
 203 pavement model, the acquisition avoids the PML (Perfectly Matched Layer) zone of
 204 the model. In addition, 25 A-scans on both sides of the pavement model are ignored to
 205 eliminate the A-scans with edge effects.
 206 A total of 150 A-scans are acquired to create a B-scan consisting on 50 A-scans for
 207 debonding case and 100 A-scans for non-debonding case. The B-scans are generated
 208 to approximate the two radar configurations used in the experiments.

209 3. OCSVM for Thin debonding detection

210 In this section, we present the different preprocessing techniques and the detection
 211 process based on the proposed method for rapid short-term debonding detection. Short-
 212 term detection refers to the debonding detection spanning over various time intervals
 213 carried out without prior knowledge of the initial conditions of the pavement. This ap-
 214 proach enables the operator to efficiently locate debondings (outliers) even from large
 215 GPR datasets.

216 *3.1. Basics on GPR time signatures*

217 For our application, the top two pavement layers are studied. As mentioned in
218 (Todkar et al., 2019), in case of a healthy pavement, two reflected echoes are obtained:
219 backscattered echo from the top surface of the pavement and echo from the interface
220 between the first and the second pavement layers. The signal amplitude is proportional
221 to the contrast in the dielectric of the two media (pavement layers).

222 In case of a debonded pavement, three echoes are expected. The first is the backscat-
223 tered surface echo while the two subsequent echoes are reflections from the upper and
224 the lower interfaces of the debonded layer that overlap in time domain. This overlap-
225 ping is mostly constructive as long as the debonding thickness is limited by the quarter
226 wavelength ($th_{deb} \leq \lambda_{deb}/4$) (Todkar et al., 2019). Since the amplitude of the reflected
227 echo depends on the dielectric contrast at each interface of the stratified medium, the
228 first (or surface) echo is the strongest whereas the second strongest echo is seen at
229 the interface (healthy and debonded zones) and can be up to $\frac{1}{10}$ th of the surface echo.
230 Then, the peak magnitude of the second strongest echo is used to center the time win-
231 dow. Therefore, the echoes can be identified based on both the amplitude and time
232 delays.

233 *3.2. Pre-processing GPR data*

234 Signal features can be generally defined as the characteristics of a data set that
235 contains sufficient information to distinguish among the classes of data (Liu and Mo-
236 toda, 1998; Siegwart et al., 2011), which, in our application are debonding and non-
237 debonding classes.

238 The useful information expected from the signal features are that they (a) retain the
239 characteristics of the original data (b) are limited in number, and, (c) are unaffected by
240 unwanted variability in the input data (Kumar and Bhatia, 2014). The study of the use
241 of signal features in this paper is necessary as it helps to:

- 242 • validate the adaptability of OCSVM for various input data types (*i.e.*, Sensitivity
243 Analysis)
- 244 • reduce the processing data size and complexity
- 245 • avoid over-learning/over-fitting that may hinder the performance, and,
- 246 • possibly improve the debonding detection rate

Signal features can either be global or local depending on the zone from where they are extracted. (Peeters, 2004) defines global features as the characteristics computed from the complete signal and local features as the characteristics computed within a specific time window.

For our application, we define the global and local features by the size of the time-gating window. The global features are computed within the time window that encompasses the surface and the interface echo, which, as shown in Figure 4 for experimental data is approximately between 1.84 ns and 4.7 ns. The time-gating window to compute local features on the other hand, is limited at the interface (*i.e.*, echo between the top

and base pavement layers). The length of time gating window (W_{tg}), in terms of number of samples for local features is a function of the center frequency of the GPR pulse given by (Todkar et al., 2019):

$$W_{tg} = 2 \times \{f_s \times t_w\} \quad (1)$$

247 where, f_s is the sampling frequency and t_w is the pulse width of the emitted GPR
 248 pulse.

A tapered Tukey-cosine window is used for the interval W_{tg} in order to extract local

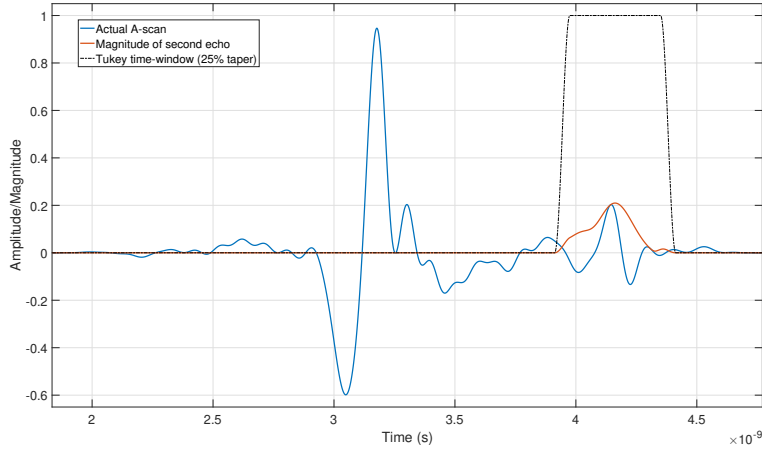


Figure 4: Time-gating window for an A-scan from APT data (Tack-free based defects at 10K loading) (Todkar, 2019)

249 features. For better representation, Figure 4 shows an A-scan with the window. Fig-
 250 ure 5 compares a B-scan before and after the time gating is performed.

251 Following (Todkar et al., 2019), the local signal features are used in this paper to
 252 achieve better debonding detection.

253 For the sake of comparison, this paper uses the same local signal features listed in Ta-
 254 ble I of (Todkar et al., 2019). This local feature set consists of 10 signal features for
 255 each A-scan data vector. These features are chosen using the Ground truth (GT) as-
 256 sociated to a known GPR dataset to perform the feature selection. GT categorizes the
 257 A-scans into two classes, namely debonding and non-debonding cases. The probability
 258 density function (PDF) of the features is computed over each class to determine those
 259 which are sensitive enough to the pavement conditions. The separation between the
 260 two curves is used to choose the best features for our application (Todkar et al., 2019).
 261 In order to use the OCSVM method, the feature matrix (\mathbf{F}_{mat}) is column-normalized
 262 (A-scan by A-scan) between $[-1, 1]$ using Min-Max normalization. By normalization,
 263 the impact of large valued features extracted on a different scale is reduced allowing
 264 small valued features to contribute equally.
 265

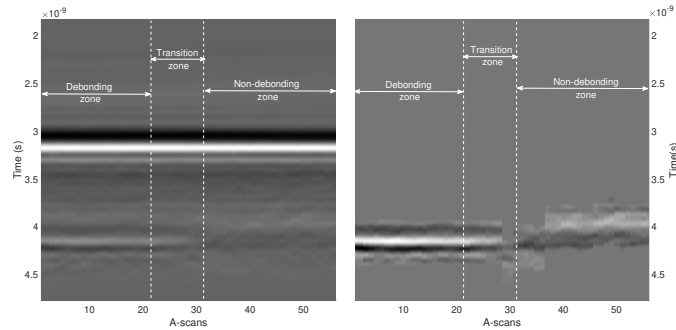


Figure 5: Comparison of the original and the time-gated B-scans from APT data (Tack-free based defects at 10K loading) (Todkar, 2019). The amplitudes in both images have been normalized for visualization

266 **3.3. Proposed debonding detection method**

(Schölkopf et al., 2001) introduced the OCSVM as a method of adapting SVM to a one-class classification problem.

For OCSVM, the learning data consists only in non-debonding data and the origin is the only member of the debonding data set (by contrast, SVM would require the learning data set to include both debonding and non-debonding data). Figure 6 shows the geometrical interpretation of OCSVM.

OCSVM can be obtained using Support Vector Data Description (SVDD) wherein the

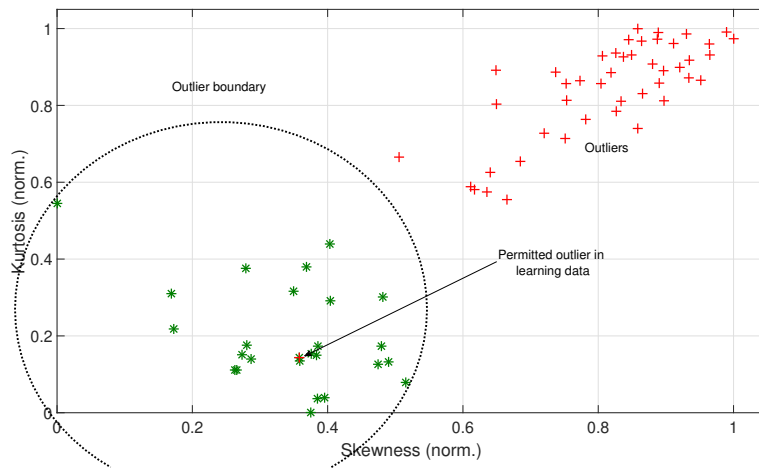


Figure 6: Geometrical representation of a OCSVM using two signal features (kurtosis and skewness) from APT data (Sand-based defects)

data-points are mapped onto a higher dimension circumscribed by a hyper-sphere (Tax and Duin, 2004). This hyper-sphere acts as the limiting factor and the aim would be to

find this hyper-sphere with minimum radius that encompasses the positive (*i.e.* Non-debonding) data mapped on the said hyper-space.

To begin with, we first define a closed hyper-sphere around the known non-debonding learning feature set in the hyper-space. This sphere is defined by its radius $R > 0$ and centered at b . The learning set strictly contains the data-vectors of only one type of data and is given by: $\mathbf{X}_{mat} = \mathbf{x}_1, \mathbf{x}_2 \dots, \mathbf{x}_N$ where N is the number of learning data-vectors and $\mathbf{X}_{mat} \in \mathbf{F}_{mat}$ is the feature matrix of the known non-debonding zone.

We aim to minimize the volume of this hyper-sphere by minimizing the radius R keeping in mind that all the learning points are embedded within this hyper-sphere. To do so, as (Tax and Duin, 2004) mentions, an error function is defined analogous to Support vector classifiers as:

$$f(R, b) = R^2$$

$$\text{constrained to: } \|\mathbf{x}_i - b\|^2 \leq R^2 \quad \forall i \quad (2)$$

Now, to allow all possibilities of the outliers into the learning set, the distance between each data-vector \mathbf{x}_i and the center of the hyper-sphere b should be less than R^2 . If this condition is not satisfied, the distances larger than this should be penalized. To do so, a positive slack variable ξ_i (where $\xi_i \geq 0$) is introduced. The minimization problem thus becomes (Chen et al., 2001; Tax and Duin, 2004):

$$\text{minimize}_{R, \xi, b} f(R, b, \xi) = R^2 + \frac{1}{\nu N} \sum_{i=1}^N \xi_i \quad (3)$$

$$\text{s.t. } \|\mathbf{x}_i - b\|^2 \leq R^2 + \xi_i, \quad \forall i = 1, 2, \dots, N$$

267 where,

- 268 - ν is a user designed parameter that is used to determine the amount of
- 269 admitted slack (*i.e.* amount of outliers permitted in the learning data)
- 270 - ξ_i is the introduced slack variable

271 Equation 3 is the standard form of convex quadratic programming problem which can
 272 be solved by introducing Lagrangian multipliers, α_i and β_i (Gu et al., 2010; Tax and
 273 Duin, 2004):

$$\mathcal{L}(R, b, \alpha, \beta, \xi) = R^2 + \frac{1}{\nu N} \sum_{i=1}^N \xi_i + \sum_{i=1}^N \alpha_i \left[R^2 + \xi_i - \|\mathbf{x}_i - b\|^2 \right] + \sum_{i=1}^N \beta_i \xi_i \quad (4)$$

By solving the Lagrangian and substituting values, we obtain:

$$\mathcal{L}(R, b, \alpha, \beta, \xi) = \sum_i \alpha_i \langle \mathbf{x}_i, \mathbf{x}_i \rangle - \sum_i \sum_j \alpha_i \alpha_j \langle \mathbf{x}_i, \mathbf{x}_j \rangle \quad (5)$$

274 The $\langle \cdot, \cdot \rangle$ dot product can be replaced with a kernel function ($\Phi(\cdot)$) in One-class
 275 SVM (Gu et al., 2010).

276 For cases where the data is linearly separable, $\Phi(\cdot)$ is a linear kernel, otherwise, a non-
 277 linear kernel (namely, Gaussian radial basis function, sigmoid or polynomial kernel
 278 (Chang and Lin, 2011)) is used. The boundary condition is then defined by R as:

$$\|\Phi(\mathbf{x}_i) - b\|^2 = R^2 \quad (6)$$

279 where,

- 280 - \mathbf{x}_i is any support vector with $0 < \alpha_i < \frac{1}{\nu N}$
- 281 - α_i is the Lagrange multiplier

The OCSVM decision function $D(\mathbf{x})$ is given as:

$$D(\mathbf{x}) = \text{sgn}\{\Phi(\mathbf{x}_i) - b\} = \begin{cases} +1, & \text{for } f(x) \\ -1, & \text{for } \bar{f}(x) \end{cases}$$

282 where, $f(x)$ and $\bar{f}(x)$ are respectively inlier and outlier data predicted by OCSVM.

283 3.4. OCSVM programming

284 The anomaly debonding detection is implemented in Python Scikit-learn (Pedregosa
285 et al., 2011) using the LIBSVM (Chang and Lin, 2011) library to create a outlier detec-
286 tion model. The latter model requires to select the kernel function in Equation 3 and its
287 associated parameters according to the processing presented in the next section. The
288 selected model is then used to identify the presence or absence of debondings (outliers)
289 over the unknown pavement area.

290 The learning data size for synthetic GprMax and experimental data are respectively 30
291 and 20 adjacent A-scans from the healthy pavement zone. In case of simulated dataset,
292 the learning data is considered separately for each individual noise level and debonding
293 thickness. For experimental data, the learning data for each B-scan is acquired from
294 the healthy pavement zone at their respective loading cycle of the fatigue carousel.

295 The outlier-detection model is generated using the learning data. A k -fold Cross
296 validation technique (with $k = 3$) is used alongside a loss function to determine the
297 model parameters. These parameters are then used to identify the presence or absence
298 of debondings over the unknown test pavement area. Figure 7 briefly depicts the over-
299 all steps carried out to detect thin debondings using OCSVM.

300 Since learning data consists of only one type of data, it is expected to demonstrate
301 similar performance levels at various values of learning data set size *i.e.*, N_{learn} . Nev-
302 ertheless, this analysis will be studied in the next section (Section 4.2.2).

303 3.5. Choice of OCSVM parameters

The selection of the kernel function and the associated parameters is conducted by
optimizing the classification results over the training data set. In this paper, two kernel
functions namely, linear and non-linear (Gaussian) have been compared on this basis.
These kernel functions can be expressed as (Eude and Chang, 2018):

$$\Phi_{lin}(\mathbf{x}_i, \mathbf{x}_j) = \mathbf{x}_i^T \mathbf{x}_j + C \quad (7)$$

$$\Phi_{rbf}(\mathbf{x}_i, \mathbf{x}_j) = \exp\{-\gamma\|\mathbf{x}_i - \mathbf{x}_j\|^2\} \quad (8)$$

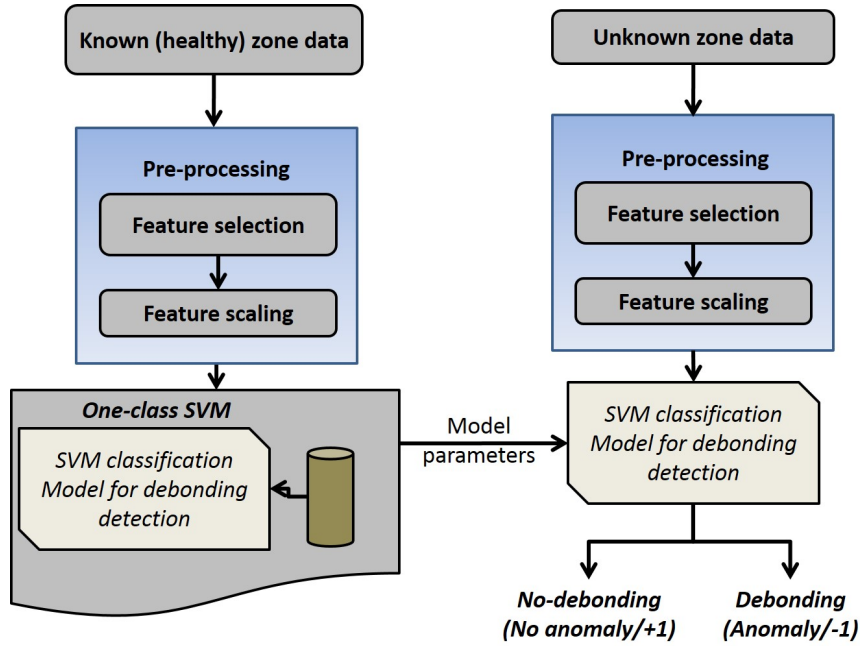


Figure 7: Debonding detection process (Todkar, 2019)

304 The linear kernel requires the optimization of a single parameter, namely, ν ; whereas
 305 the Gaussian kernel uses both ν and γ parameters.

306 The Hinge loss is usually the preferred loss function since it is better suited for classifi-
 307 cation problems; by contrast, the root-mean square error would be preferred instead for
 308 regression problems (Todkar, 2019). For our intended output, the Hinge loss function
 309 only takes two discrete values $[0, 2]$, *i.e.* non-zero values for misclassified A-scans
 310 (False Positive - FP and False Negative - FN values) and to zero values for properly
 311 classified A-scans (True Positive - TP value).

312 In order to gain further sensitivity in the optimization problem, the average Hinge-loss
 313 function is computed at each signal-to-noise ratios (SNR) over the 100 independent
 314 noise realizations, accordingly to the Monte-Carlo process. As a result, the average
 315 hinge loss function then takes intermediate step-wise value within the $[0, 2]$ interval,
 316 as traced in both Figure 8 and Figure 9.

317 The search for the minimum hinge loss value is usually performed over a fine grid
 318 search in ν and γ . The user designed slack is limited to $\nu \in]0, 1]$. In practice, the
 319 ν -parameter decides the fraction of outliers present in the training data set (Mei and
 320 Zhu, 2015). Since the training dataset for simulated data is devoid of outliers, the value
 321 of ν is fixed to a low value, *i.e.*, $\nu = 0.001$. The optimum value is then expected to
 322 be small. On the other hand, the γ -parameter range is chosen from (Pedregosa et al.,
 323 2011), where the suitably sufficient range for γ is $[10^{-3}, 10^3]$.

324 Figure 8 presents the average hinge loss for the Gaussian kernel after 100 Monte-Carlo
 325 realizations at 30 dB SNR with 20 learning samples. As mentioned in the previous

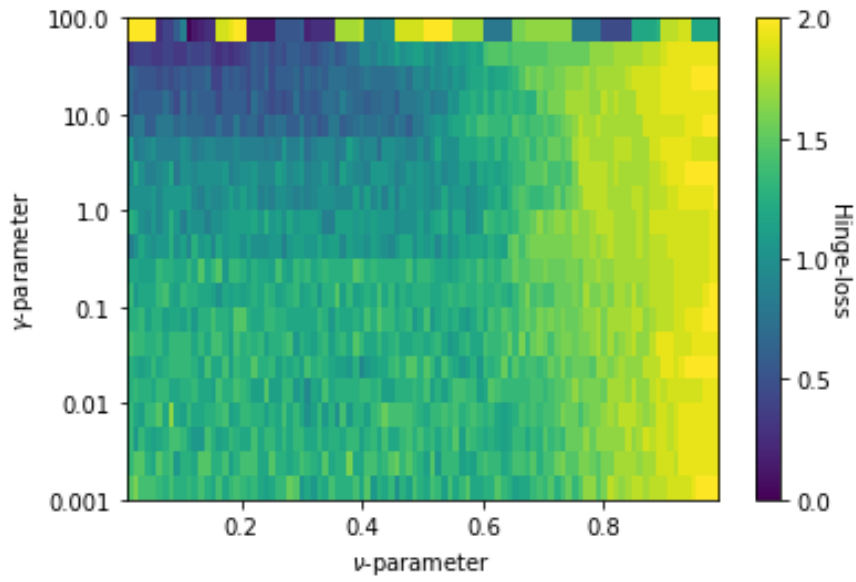


Figure 8: Variation of Hinge-loss function with respect to ν and γ parameters for a simulated gprMax B-scan image with $SNR = 30$ dB using 20 learning samples and $th_{deb} = 3$ mm

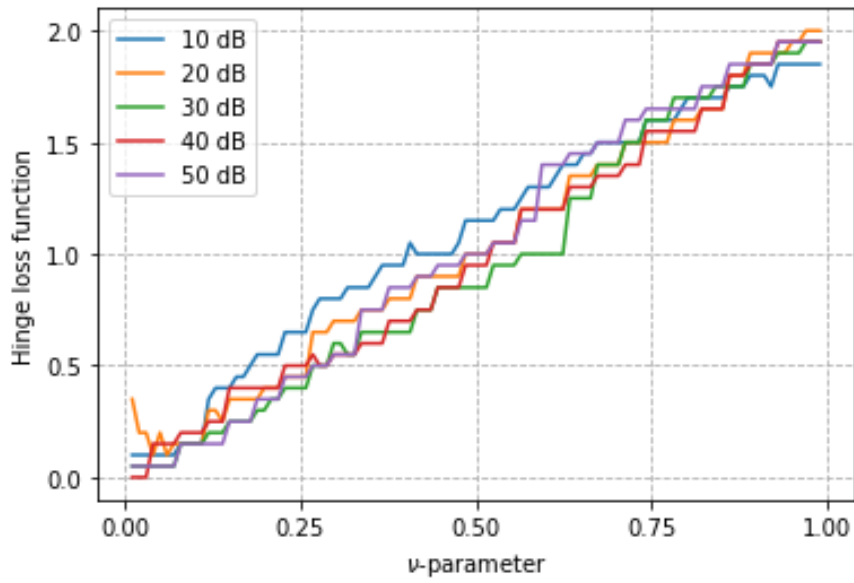


Figure 9: Variation of Hinge-loss function with respect to ν -parameter for a simulated gprMax B-scan image at various SNR levels using 20 learning samples and $th_{deb} = 3$ mm

326 section, a 3-fold cross validation is used here. It was observed that the function to min-
327 imize shows smooth step-wise variations with a global minimum loss function value
328 observed at smaller ν and intermediate γ values. It is also shown that γ parameter has a
329 small influence on the loss function at medium to high SNR levels. Further simulations
330 have shown that, at lower SNR values, the Hinge-loss variations depends on both ker-
331 nel parameters at once again smaller ν values and intermediate and/or high γ values.
332 Due to the small influence on gamma-parameter, the linear kernel function has been
333 preferred as kernel function for data with SNR = 30 dB or above, thereby reducing the
334 computation to simply one parameter, namely, ν . Figure 9 shows the influence of ν
335 over Hinge-loss values at various SNR levels. It can be confirmed that the optimum ν
336 value is close to 0.001 and slightly depends on SNR. By contrast, the Gaussian kernel
337 would be better suited for data of SNR = 20 dB or below.

338 3.6. Metrics to assess OCSVM performance

339 A benchmark, called as Ground Truth (GT) is established for each dataset. GT
340 refers to the accurate interpretation and labeling of the pavement conditions under con-
341 trolled tests. While the GT is used for the synthetic GprMax data, the experimental
342 data uses the Pseudo Ground Truth (PGT), which is a roughly estimated benchmark
343 of pavement conditions. In both cases, the debonding A-scan is assigned to 1 and the
344 non-debonding A-scan is assigned to -1.

345 The OCSVM detection can result in one-out-of-four possible outcomes namely: True
346 Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN). These
347 four values are collectively grouped as a Confusion Matrix (Kulkarni et al., 2020).
348 This matrix is used to compute the Dice score (DSC) that describes the performance
349 of OCSVM model.

350 Furthermore, for the sake of reliability, the transition zone between debonding and
351 healthy zones is not included in the performance assessment for both simulated and
352 experimental data.

353 4. Tests and validation of GprMax model

354 In this section we analyze the performance of the detection method on pavement
355 model generated using GprMax.

356 4.1. Noiseless simulated data

357 Figure 10a and Figure 10b illustrate the detection results of the OCSVM method for
358 the two radar technologies. The classification labels in the figures use -1 to indicated
359 Non-debonding and +1 for debonding labels.

360 Due to the absence amount of variation in the feature distribution for the healthy zone
361 (non-debonding A-scans along the scanning direction), a small data size ($N_{learn} = 20$)
362 is sufficient for the learning step. Due to the ideal smooth interface B-scan data, it was
363 observed that OCSVM presented an excellent detection result with zero false detection.
364 As mentioned in Section 3.5, a linear kernel was used and the value for ν parameter
365 was set to 0.001.

366 However, in order to explore the OCSVM capability to adapt to various pavement
367 configurations and scenarios, Sensitivity analysis is presented hereafter.

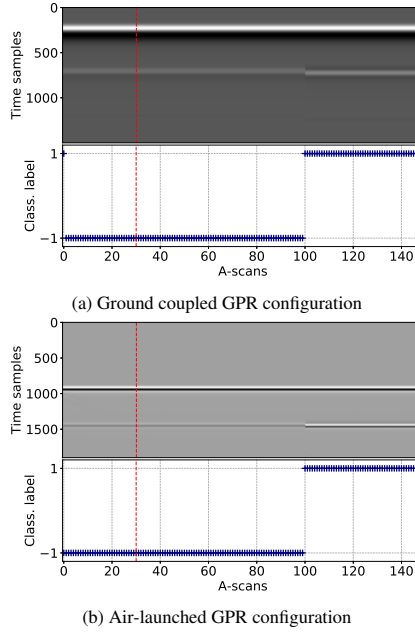


Figure 10: OCSVM debonding detection on noiseless GprMax data with debonding thickness $th_{deb} = 3$ mm. The vertical dashed line indicates the learning data

368 4.2. Sensitivity Analysis

369 Sensitivity analysis (SA) is the study of uncertainties between a given input and its
 370 expected output (Samadzadegan et al., 2010). Using SA, we can also determine the
 371 robustness and adaptability of OCSVM over various GPR technologies and pavement
 372 configurations. In this section, we present the OCSVM-SA with respect to, noise level
 373 and learning data size.

374 4.2.1. Effect of Noise

Noise plays an important role in the consistency of the signal and as such, the noise
 level in the signal sometimes limits the performance of the detection methods. The
 noise level is decided by the standard deviation of the noise given by:

$$\sigma_N = \frac{M}{10^{\left(\frac{SNR}{20}\right)}} \quad (9)$$

375 where,

- 376 - M is the maximum magnitude of the second echo of the signal (A-scan),
- 377 - σ_N is the standard deviation of the noise added,
- 378 - SNR is the signal-to-noise ratio in dB.

379 Using Equation 9, an ideal-White Gaussian noise is added to each B-scan image such
 380 that the noise samples are decorrelated to each other in both vertical (*i.e.*, time) and

381 horizontal (*i.e.*, between A-scans); the SNR has been varied between 10 dB to 60 dB.
 382 The detection results presented are averaged over 100 independent noise realizations
 383 for each Ground-coupled and Air-launched GPR configurations.
 384

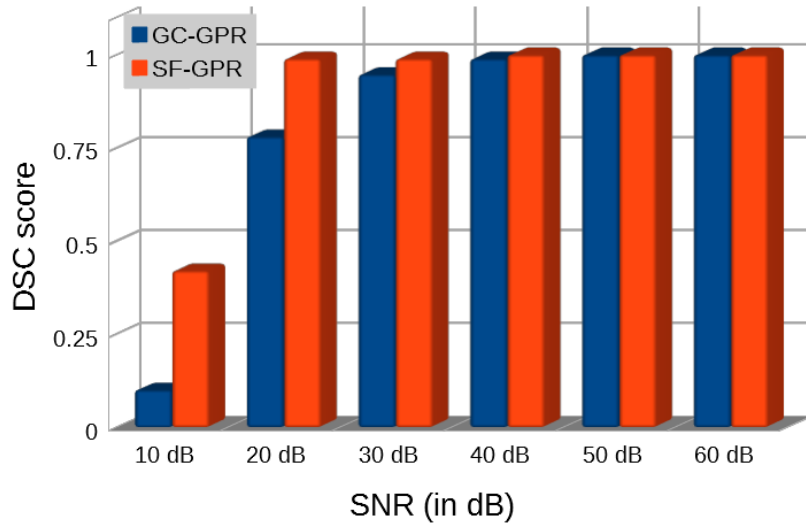


Figure 11: Comparison of average DSC score for GC-GPR and SF-GPR GprMax configurations at various levels of SNR for 100 independent Monte-Carlo realizations ($th_{deb} = 3$ mm, $N_{learn} = 30$)

385 Figure 11 traces the DSC score with respect to SNR for both GC-GPR and SF-
 386 GPR configurations. For the two GPR configurations, it is observed that at SNR =
 387 10 dB, the debonding detection indulges a high false-detection rate (both FP and FN)
 388 leading to reduced performance.

389 However, the performance rapidly improves above 0.75 beyond SNR = 20 dB. In
 390 all cases, the detection rate is higher for SF-GPR data compared to GC-GPR. The
 391 higher frequency and larger frequency bandwidth provide better time resolution to SFR
 392 radar, that provides better separation of surface and interface echoes. This conclusion
 393 is worth over all levels of SNR according to Figure 11, but also for any thickness as
 394 well according to (Todkar, 2019).

395 During this analysis, the learning data size was chosen by intuition. However, the
 396 learning data has an effect over the performance of SVM as stated in (Huang, 2001;
 397 Thomas et al., 2015). Thus, in the following section, the sensitivity of the OCSVM is
 398 tested with respect to, the learning data size and its effect on the debonding detection
 399 rate.

400 4.2.2. Effect of learning data size

401 For Two class SVM, learning to test data ratio used in the literature varies from as
 402 low as $\frac{1}{5}$ th (Huang, 2001) to as high as $\frac{4}{5}$ th (Thomas et al., 2015).
 403 On the other hand, in case of OCSVM, the literature suggests to use $\frac{2}{3}$ rd of the data for

404 learning (Dreiseitl et al., 2010; Li et al., 2003). Since One-class SVM relies on only
 405 one type of data to create the classification model, the dependability of the learning
 406 data size is inclined on the variations of the non-debonding data.
 407 As specified in Section 4.1, the noiseless data presents no variation in the feature dis-
 408 tribution. Thus the learning data size can be very small. However for noisy data, the
 409 learning data size is expected to play a vital role in the debonding detection. Therefore,
 410 in this section, we analyze the effect of learning data size on the performance.
 411 Here, the debonding thickness is set to 3 mm and noise is added to obtain an SNR of
 412 30 dB. Figure 12 traces the DSC score variation at increasing learning data sizes for
 413 both GC-GPR and SF-GPR configurations. These are obtained by a Monte-Carlo pro-
 414 cess averaged over 100 independent realizations for the two configurations.
 It can be observed that, for both GC-GPR and SF-GPR configurations, for small learn-

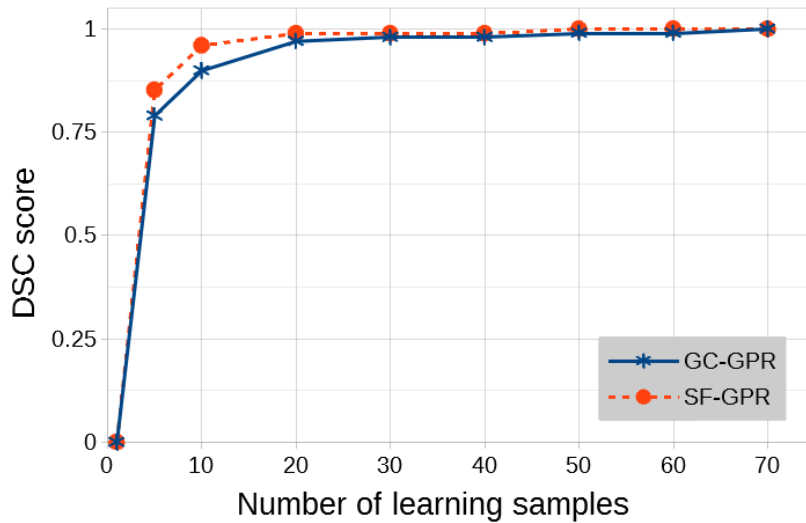


Figure 12: DSC variation with respect to learning data size for noisy GprMax data with SNR = 30 dB for ground-coupled and air-launched configurations ($th_{deb} = 3$ mm, $N_{learn} = 30$)

415 ing datasets (= 5 A-scans), the debonding detection rate (DSC) is greater than 0.8.
 416 However a rapid increase in the performance (*i.e.*, DSC) is observed up to $N_{learn} =$
 417 20 beyond which a gradual increase in DSC up to 1 was observed. This SA infers that
 418 the OCSVM method can be implemented on GPR datasets for debonding detection
 419 irrespective of its configuration.
 420

421 4.2.3. Effect of debonding thickness

422 In the previous sections, we studied the effect of various parameters such as noise,
 423 learning data size *etc.* on the performance. However, in every case, the debonding
 424 layer within the B-scan was fixed to one thickness (*i.e.*, 3 mm). In this section, we
 425 analyze the performance behavior in conditions with different debonding thicknesses.

426 As done in previous sections, the detection is averaged over 100 independent Monte-
427 Carlo realizations.

428 A B-scan is generated with the pavement parameters specified in Section 2.2. The
429 defective zone consists of four sub-zones containing debondings of respectively 1 mm,
430 3 mm, 5 mm and 7 mm thicknesses (as depicted in Figure 13). Each defective zone is
431 0.2 m long with 20 acquired A-scans whereas the non-defective zone is 1 m long with
432 100 acquired A-scans. Thus, the resulting B-scan contains a total of 180 A-scans.

Figure 14a and Figure 14b respectively present the debonding detection for GC-GPR

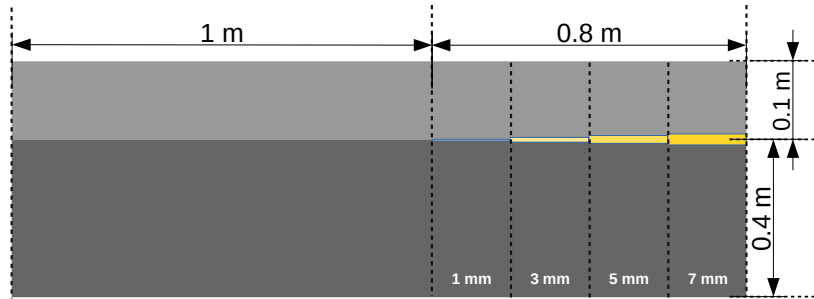


Figure 13: Pavement structure model used to generate a B-scan with defective zone containing debondings of respectively 1 mm, 3 mm, 5 mm and 7 mm thicknesses. Antennas are positioned respectively as shown in Figure 3

433 and SF-GPR configurations of the pavement model from Figure 13. For both SF-GPR
434 and GC-GPR configurations, zero non-detection (*i.e.*, FN) were observed. However,
435 some false alarms were seen to occur in both cases leading to the DSC scores of 0.98
436 and 0.99 for respectively GC-GPR and SF-GPR configurations.

437 Since the SF-GPR operates at a higher frequency, the time resolution of its data is
438 higher. This provides a better feature separation and therefore has a better performance
439 than the GC-GPR data.

440 In the final section, we implement One-class SVM debonding detection method on
441 experimental data collected at the IFSTTAR's APT site using both air-launched and
442 ground-coupled radar configurations.
443

444 5. Experimental validation of OCSVM

445 In this section, the proposed method is tested on the experimental data presented in
446 Section 5. Here, the OCSVM model optimization and the debonding detection process
447 is the same as done for GprMax pavement model in Section 4. Two pavement sections
448 are probed with the GPR. The first is the known pavement section with no debondings
449 consisting of $N = 20$ A-scans and is used for the 'learning' step (as done in Section 4).
450 The second section is then probed whose subsurface condition is unknown. This un-
451 known B-scan is divided by the operator into two groups namely, transition zone and
452 testing zone. The transition zone is the region where the presence/absence of debond-

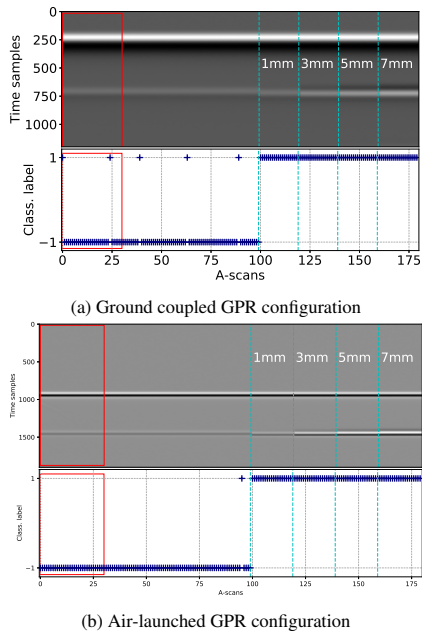


Figure 14: One-class SVM debonding detection on noisy GprMax data (SNR = 30 dB, $N_{learn} = 30$) with various debonding thicknesses $th_{deb} = 1$ mm, 3 mm, 5 mm and 7 mm. The solid box indicates the learning zones data; vertical dashed lines indicate various thickness zones

453 ing is ambiguous and this zone is excluded from the performance evaluation of the
 454 detection method.

Loading stages	SFR	GSSI
10K	0.96	0.93
50K	0.97	0.9
100K	0.96	0.96

Table 1: Debonding detection DSC score for Tack-free defects at various loading stages for SFR and GSSI-GPR data

455 Table 1 presents DSC score for Tack-free defects each at various loading stages.
 456 Figures 15, 16 and 17 present the detection results for Tack-free defects at 10K, 50K
 457 and 100K cycles loading stages. The first $N = 20$ A-scans is the known region used
 458 for the learning step. In each figure, the region between the dashed lines is the transi-
 459 tion zone. The remainder is the testing B-scan data.

460 In case of experimental data, the thickness of Tack-free defects was observed to be be-
 461 tween 2 mm to 3 mm (see Section 2.1). The OCSVM method provided similar results
 462 for GC-GPR and SF-GPR data at 10K and 100K loading. However the performance in
 463 case of GC-GPR data at 50K loading was reduced due to aberrant A-scans present in

464 the learning data.

465 From Figures 15, 16 and 17 and Table 1, it can be observed that OCSVM is capable
466 of detecting debondings with high efficiency and very few false alarms. These false
467 alarms could be attributed to the data-points present very close to the OCSVM hyper-
468 sphere boundary. The zone in the blue box indicates the learning data and t_{zone} denotes
the transition zone. In general, it was observed that the performance of the OCSVM

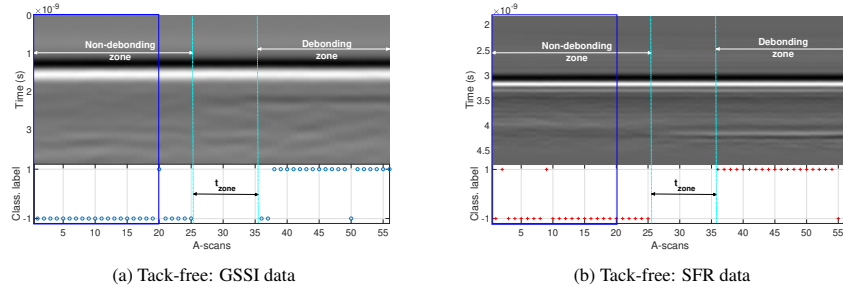


Figure 15: Debonding detection at 10K loading for for Tack-free based defects

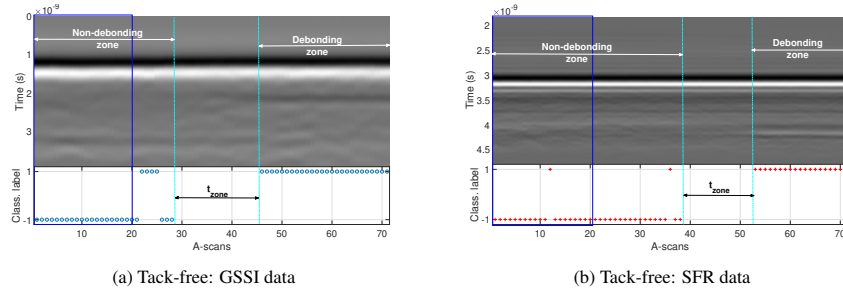


Figure 16: Debonding detection at 50K loading for for Tack-free based defects

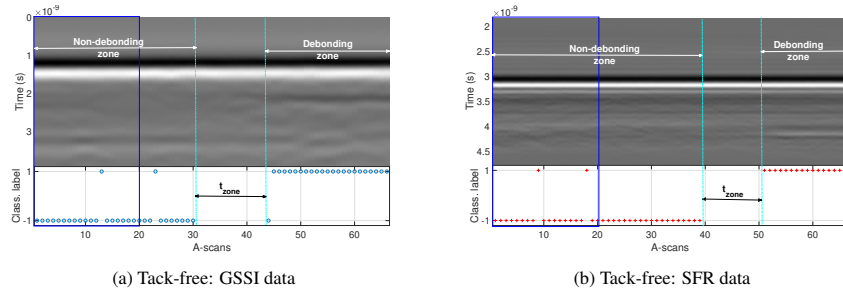


Figure 17: Debonding detection at 100K loading for for Tack-free based defects

469 method was always slightly better for SF-GPR than the GC-GPR data. This observation
470 was held true for both synthetic GprMax as well as experimental data. The SF-GPR
471

operates at a higher frequency ($f_{c,SFR} = 4.2$ GHz) compared to the GC-GPR operating at $f_{c,GSSI} = 2.6$ GHz. The higher temporal resolution allows the SF-GPR to provide better distinction between the interface echoes and therefore, better separated features of the two classes.

However, the OCSVM method has a few limitations. In this paper, a known non-debonding dataset is used to train the OCSVM model. The presence of debonding data (outliers) in this learning dataset could skew the classification model leading to false detection (False Negative and/or False Positive). OCSVM is also limited by the number of classes for classification (inliers or outliers); this method cannot be used for multi-class classification.

6. Comparison between One-class and Two-class SVMs

In this section, we compare the performance of One-class and Two-class SVM (from (Todkar et al., 2019)). Contrary to OCSVM, the Two-class SVM requires data from both classes during the learning stage. The best separation in this case is intuitively achieved with the hyper-plane which provides the maximum distance from the nearest data-points of the debonding and non-debonding classes.

In case of Two-class SVM, the initial data set (*i.e.*, 10K loading) for each defect was used in the learning stage to generate a classification model. This approach is suitable for a long-term detection while the initial conditions is available. OCSVM in this paper is more of an immediate/short-term detection approach because it does not require an initial data set. Detection is carried out on each loading stage independent of the previous data. Since (Todkar et al., 2019) performs detection on SF-GPR data only, we compare the same with OCSVM results. For the sake of relevance, we cover the two extreme cases and the results for strong (Geotextile) defects are presented in addition to Tack-free defect data.

The Two-class SVM detection (as presented in (Todkar et al., 2019)) for SF-GPR data uses approximately 50% of its data as learning data. OCSVM, on the contrary, uses a smaller learning data (first 20 A-scans, which is about 20% - 25%) to provide very similar results. Table 2 and Table 3 compare the *DSC* scores of One-class and Two-class SVMs for Tack-free and Geotextile defects respectively. It can be noted that in case of Tack-free defects, both One and Two-class SVMs provide similar performance.

Both One and Two-class SVMs demonstrate good performance. However, OCSVM

Loading stage	Two-class SVM	One-class SVM
10K	1.0 [†]	0.96
50K	0.96	0.97
100K	0.95	0.96

Table 2: Comparison of *DSC* score for Tack-free defects at various loading stages obtained using Two-class and One-class SVMs (Two-class detection results taken from (Todkar et al., 2019))

Loading stage	Two-class SVM	One-class SVM
10K	1.0 [†]	0.96
50K	1.0	1.0
100K	1.0	0.95

Table 3: Comparison of DSC score for Geotextile defects at various loading stages obtained using Two-class and One-class SVMs (Two-class detection results taken from (Todkar et al., 2019))

503 relies on the non-debonding data which is readily available contrary to the two-class
504 and a smaller learning dataset size. In addition, since OCSVM does not require initial
505 data, it is well-suited to be used in field detection for short-term practical experiments.
506

507 7. Conclusion

508 In this article, we proposed to implement OCSVM for detecting thin debondings
509 from both air-coupled and ground-coupled GPR data. The proposed approach was first
510 tested with simulated pavement models from GprMax. To study its capability to adapt
511 to various data configurations, OCSVM Sensitivity Analysis was conducted.

512 With the help of OCSVM, we were able to detect debondings with $DSC > 0.75$ (for
513 GC-GPR) and $DSC > 0.98$ (for SF-GPR) at an $SNR = 20$ dB or greater. At a small
514 learning data size of $N_{learn} = 5$, GC-GPR showed a $DSC > 0.8$ in contrast to the
515 SF-GPR with $DSC > 0.85$. The performance reached close to $DSC \approx 1$ as the learn-
516 ing data size was gradually increased. In addition, OCSVM was able to detect thin
517 debondings within a B-scan varying from 1 mm to 7 mm from both GC and SF-GPR
518 with very few false detection. Overall, for simulated data, the OCSVM method per-
519 formed better with SF-GPR data sets than its counterpart GC-GPR data sets.

520 The OCSVM was then validated on experimental data collected with two GPR config-
521 urations over embedded Tack-free defects. Over three loading stages (10K, 50K and
522 100K), due to its increased temporal resolution and better separated signals in the time
523 domain, OCSVM provided better results for SF-GPR compared to the GC-GPR data
524 sets.

525 Finally, a short comparison to the previous works showed that the OCSVM method
526 provided similar results as its Two-class counterpart. It was observed that OCSVM is
527 suited for short-term debonding detection experiments.

528 In perspective, we aim to extend the performance assessment of the proposed detection
529 method to the entire data set collected over the tested pavement structure during its
530 entire lifetime, *i.e.*, up to 800K loading.

[†](Todkar et al., 2019) used 10K data set as learning and the remaining loading sets as test; contrast to the OCSVM in this paper that uses the first 20 A-scans of each loading for learning

531 **Acknowledgment**

532 The French National Research Agency (ANR) is gratefully acknowledged for sup-
533 porting the ANR-ACIMP project (Project code: ANR-18-CE22-0020).

534 **Appendix A. Computation of SNR from experimental data**

In this appendix, we present, in brief, the estimation of SNR from experimental SF-GPR data.

Revisiting Equation 9 for SNR computation, we have:

$$SNR = 20 * \log \left\{ \frac{M}{\sigma_N} \right\} \quad (A.1)$$

535 Now, in order to compute the SNR from experimental data, two values, namely M
536 and σ_N are required. The value for M is obtained from the known learning data (non-
537 debonding zone) using the automatic time-gating window (Section 3.2). The standard
538 deviation of the noise (σ_N) is computed from the late arrival signals of the subsequent
539 layers as shown in Figure A.18.

To summarize, Table A.4 provides the SNR levels for all loading stages obtained for

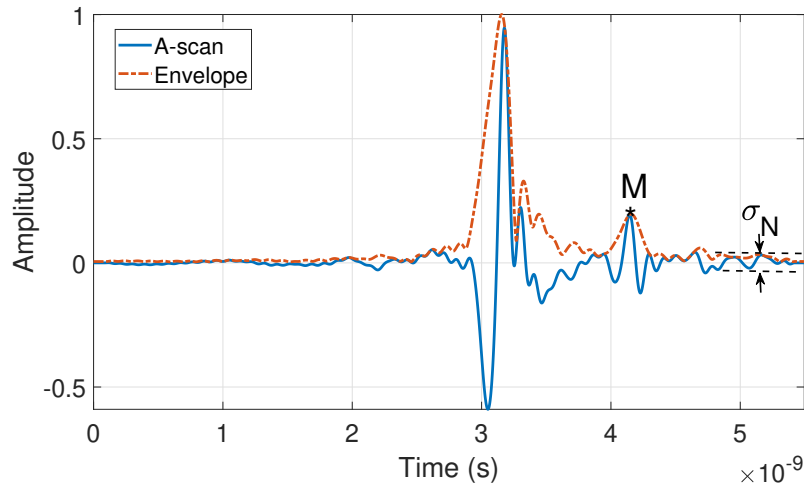


Figure A.18: Illustration of the SNR computation for GPR data from Equation 9 over the Tack-free defect at 10K loading

540

541 Tack-free defects.

542 **References**

543 Benedetto, A., Tosti, F., Ciampoli, L.B., DAmico, F., 2017. An overview of ground-
544 penetrating radar signal processing techniques for road inspections. *Signal Process-*
545 *ing* 132, 201 – 209. doi:10.1016/j.sigpro.2016.05.016.

Loading stage	SNR (dB)
10K	22.8
50K	21.9
100K	22.3

Table A.4: Estimated average SNR (in dB) of the debonded interface echo over Tack-free area at all loading stages

- 546 Büyüköztürk, O., Yu, T.Y., 2006. Understanding and assessment of debonding failures
547 in frp-concrete systems, in: Seventh International Congress on Advances in Civil
548 Engineering, Citeseer. pp. 11–13.
- 549 Chang, C.C., Lin, C.J., 2011. LIBSVM: A library for support vector machines. *ACM*
550 *Transactions on Intelligent Systems and Technology* 2, 27:1–27:27. Software avail-
551 able at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- 552 Chen, Y., Zhou, X.S., Huang, T.S., 2001. One-class svm for learning in image retrieval,
553 in: *Proceedings 2001 International Conference on Image Processing*, pp. 34–37.
554 doi:10.1109/ICIP.2001.958946.
- 555 Dérobert, X., Baltazart, V., Simonin, J.M., Norgeot, C., Doué, S., Durand, O., Todkar,
556 S.S., 2020. Radar database collected over artificial debonding pavement structures
557 during apt at the ifsttar’s fatigue carousel, in: Chabot, A., Hornych, P., Harvey, J.,
558 Loria-Salazar, L.G. (Eds.), *Accelerated Pavement Testing to Transport Infrastructure*
559 *Innovation*, Springer International Publishing, Cham. pp. 632–639.
- 560 Diakité, C., Fortino, N., Dauvignac, J.Y., 2015. Antenne ETSA ”Exponential Tapered
561 Slot Antenna” miniature pour radar à pénétration de surface, in: *19èmes Journées*
562 *Nationales Microondes (JNM 2015)*, Bordeaux, France. pp. Session Antennes large
563 bande et multibandes C3L–D, papier 9337.
- 564 Dong, Z., Ye, S., Gao, Y., Fang, G., Zhang, X., Xue, Z., Zhang, T., 2016. Rapid de-
565 tection methods for asphalt pavement thicknesses and defects by a vehicle-mounted
566 ground penetrating radar (gpr) system. *Sensors* 16. doi:10.3390/s16122067.
- 567 Dreiseitl, S., Osl, M., Scheibböck, C., Binder, M., 2010. Outlier detection with one-
568 class svms: An application to melanoma prognosis. *AMIA: Annual Symposium*
569 *proceedings*. *AMIA Symposium 2010*, 172–176.
- 570 Dérobert, X., Baltazart, V., 2012. Report on step-frequency radar data collection on the
571 fatigue carousel. Personal communication (2012) .
- 572 Dérobert, X., Baltazart, V., Simonin, J.M., Norgeot, C., Durand, O., Todkar, S.,
573 2020. Radar database over large debonded areas into pavement structures. doi:10 .
574 25578/N0RXSQ.

- 575 Ebrahim, S.M., Medhat, N., Mansour, K.K., Gaber, A., 2018. Examination of soil
576 effect upon gpr detectability of landmine with different orientations. *NRIAG Journal*
577 *of Astronomy and Geophysics* 7, 90 – 98. doi:10.1016/j.nrjag.2017.12.
578 004.
- 579 Eude, T., Chang, C., 2018. One-class svm for biometric authentication by keystroke
580 dynamics for remote evaluation. *Computational Intelligence* 34, 145–160. doi:10.
581 1111/coin.12122.
- 582 Giannopoulos, A., 2005. Modelling ground penetrating radar by gprmax. *Construction*
583 *and building materials* 19, 755–762.
- 584 Gu, H., Zhao, G., Qiu, J., 2010. One-class support vector machine with relative
585 comparisons. *Tsinghua Science and Technology* 15, 190–197. doi:10.1016/
586 S1007-0214(10)70049-3.
- 587 Hajj, E., Loria, L., Sebaaly, P., 2010. Performance evaluation of asphalt pavement
588 preservation activities. *Transportation Research Record* 2150, 36 – 46.
- 589 Huang, T.S., 2001. One-class svm for learning in image retrieval, in: *Proceedings 2001*
590 *International Conference on Image Processing (Cat. No.01CH37205)*, pp. 34–37
591 vol.1. doi:10.1109/ICIP.2001.958946.
- 592 Kaplanvural, I., Peksen, E., Özkap, K., 2018. Volumetric water content estimation
593 of c-30 concrete using gpr. *Construction and Building Materials* 166, 141 – 146.
594 doi:10.1016/j.conbuildmat.2018.01.132.
- 595 Kulkarni, A., Chong, D., Batarseh, F.A., 2020. 5 - foundations of data imbalance
596 and solutions for a data democracy, in: Batarseh, F.A., Yang, R. (Eds.), *Data*
597 *Democracy*. Academic Press, pp. 83 – 106. doi:https://doi.org/10.1016/
598 B978-0-12-818366-3.00005-8.
- 599 Kumar, G., Bhatia, P.K., 2014. A detailed review of feature extraction in image pro-
600 cessing systems, in: *2014 Fourth International Conference on Advanced Computing*
601 *Communication Technologies*, pp. 5–12. doi:10.1109/ACCT.2014.74.
- 602 Li, K.L., Huang, H.K., Tian, S.F., Xu, W., 2003. Improving one-class svm for
603 anomaly detection, in: *Proceedings of the 2003 International Conference on Ma-*
604 *chine Learning and Cybernetics*, IEEE Computer Society, Los Alamitos, CA, USA.
605 pp. 3077,3078,3079,3080,3081. doi:10.1109/ICMLC.2003.1260106.
- 606 Liu, H., Motoda, H., 1998. *Feature Selection Aspects*. Springer US, Boston, MA. pp.
607 43–72. doi:10.1007/978-1-4615-5689-3_3.
- 608 Mei, S., Zhu, H., 2015. A novel one-class svm based negative data sampling method
609 for reconstructing proteome-wide htlv-human protein interaction networks, in: *Sci-*
610 *entific reports*, pp. 1–13. doi:10.1038/srep08034.
- 611 Nesbitt, I., Simon, F.X., Paulin, T., 2019. readgssi: an open-source tool to read and
612 plot GSSI ground-penetrating radar data. doi:10.5281/zenodo.3352438.

- 613 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blon-
614 del, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cour-
615 napeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine
616 learning in Python. *Journal of Machine Learning Research* 12, 2825–2830.
- 617 Peeters, G., 2004. A large set of audio features for sound description (similarity and
618 classification) in the CUIDADO project. Technical Report. Icram.
- 619 Pinel, N., Le Bastard, C., Liu, L., Bourlier, C., Wang, Y., 2009. Rough thin pave-
620 ment thickness estimation by GPR, in: *International Geoscience and Remote Sensing*
621 *Symposium, Le Cap (Cape Town), South Africa*. p. 1255. 4 pages.
- 622 Salem, M.B., Stolfo, S.J., 2012. A comparison of one-class bag-of-words user be-
623 havior modeling techniques for masquerade detection. *Sec. and Commun. Netw.* 5,
624 863–872. doi:10.1002/sec.311.
- 625 Samadzadegan, F., Hasani, H., Partovi, T., 2010. Sensitivity analysis of support vector
626 machine in classification of hyperspectral imagery, in: *Proceedings of the Canadian*
627 *Geomatics Conference, Calgary, Canada*, pp. –.
- 628 Schölkopf, B., Platt, J.C., Shawe-Taylor, J.C., Smola, A.J., Williamson, R.C., 2001.
629 Estimating the support of a high-dimensional distribution. *Neural Comput.* 13,
630 1443–1471. doi:10.1162/089976601750264965.
- 631 Senf, A., Chen, X.w., Zhang, A., 2006. Comparison of one-class svm and two-class
632 svm for fold recognition, in: King, I., Wang, J., Chan, L.W., Wang, D. (Eds.),
633 *Neural Information Processing, Springer Berlin Heidelberg, Berlin, Heidelberg*. pp.
634 140–149.
- 635 Siegwart, R., Nourbakhsh, I.R., Scaramuzza, D., 2011. Introduction to Autonomous
636 Mobile Robots: Perception. MITP. pp. 101–263.
- 637 Simonin, J.M., Baltazart, V., Le Bastard, C., Dérobert, X., 2016. Progress in monitor-
638 ing the debonding within pavement structures during accelerated pavement testing
639 on the fatigue carousel, in: *8th RILEM International Conference on Mechanisms*
640 *of Cracking and Debonding in Pavements, Springer Netherlands, Dordrecht*. pp.
641 749–755.
- 642 Simonin, J.M., Fauchard, C., Hornych, P., Guilbert, V., Kerzrého, J.P., Trichet, S.,
643 2012. Detecting unbounded interface with non destructive techniques, in: Scarpas,
644 A., Kringos, N., Al-Qadi, I., A., L. (Eds.), *7th RILEM International Conference on*
645 *Cracking in Pavements, Springer Netherlands, Dordrecht*. pp. 179–190.
- 646 Tax, D.M., Duin, R.P., 2004. Support vector data description. *Machine Learning* 54,
647 45–66. doi:10.1023/B:MACH.0000008084.60811.49.
- 648 Tbarki, K., Said, S.B., Ksantini, R., Lachiri, Z., 2017a. Landmine detection im-
649 provement using one-class svm for unbalanced data, in: *2017 International Con-*
650 *ference on Advanced Technologies for Signal and Image Processing (ATSIP)*, pp.
651 1–6. doi:10.1109/ATSIP.2017.8075597.

- 652 Tbarki, K., Said, S.B., Ksantini, R., Lachiri, Z., 2017b. One-class svm for landmine de-
653 tecton and discrimination. 2017 International Conference on Control, Automation
654 and Diagnosis (ICCAD) , 309–313.
- 655 Tbarki, K., Said, S.B., Ksantini, R., Lachiri, Z., 2018. Covariance-guided landmine de-
656 tecton and discrimination using ground-penetrating radar data. *International Journal
657 of Remote Sensing* 39, 289–314. doi:10.1080/01431161.2017.1382746.
- 658 Tellez, O.L.L., Scheers, B., 2017. Ground-penetrating radar for close-in mine detec-
659 tion. *Mine Action: The Research Experience of the Royal Military Academy of
660 Belgium* , 51.
- 661 Thomas, A., Feuillard, V., Gramfort, A., 2015. Calibration of one-class svm for mv set
662 estimation, in: 2015 IEEE International Conference on Data Science and Advanced
663 Analytics (DSAA), IEEE. pp. 1–9.
- 664 Todkar, S.S., 2019. Monitoring subsurface conditions of pavement structures using
665 Ultra-wideband radar technology. Ph.D. thesis. University of Nantes, France.
- 666 Todkar, S.S., Bastard, C.L., Baltazart, V., Ihamouten, A., Dérobert, X., 2019. Perform-
667 ance assessment of svm-based classification techniques for the detection of arti-
668 ficial debondings within pavement structures from stepped-frequency a-scan radar
669 data. *NDT & E International* 107, 102128. doi:https://doi.org/10.1016/
670 j.ndteint.2019.102128.
- 671 Todkar, S.S., Bastard, C.L., Ihamouten, A., Baltazart, V., Dérobert, X., 2018. Compar-
672 ative study of classification algorithms to detect interlayer debondings within pave-
673 ment structures from step-frequency radar data, in: 2018 38th Annual IEEE Geo-
674 science and Remote sensing society(IGARSS), pp. 1–6. doi:10.1109/IWAGPR.
675 2017.7996056.
- 676 Todkar, S.S., Bastard, C.L., Ihamouten, A., Baltazart, V., Dérobert, X., Fauchard, C.,
677 Guilbert, D., Bosc, F., 2017. Detection of debondings with ground penetrating radar
678 using a machine learning method, in: 2017 9th International Workshop on Advanced
679 Ground Penetrating Radar (IWAGPR), pp. 1–6. doi:10.1109/IWAGPR.2017.
680 7996056.
- 681 Wai-Lok Lai, W., Dérobert, X., Annan, P., 2018. A review of ground penetrating
682 radar application in civil engineering: A 30-year journey from locating and testing
683 to imaging and diagnosis. *NDT & E International* 96, 58–78. doi:https://doi.
684 org/10.1016/j.ndteint.2017.04.002.
- 685 Warren, C., Giannopoulos, A., Giannakis, I., 2015. An advanced gpr modelling frame-
686 work: The next generation of gprmax, in: 2015 8th International Workshop on Ad-
687 vanced Ground Penetrating Radar (IWAGPR), pp. 1–4. doi:10.1109/IWAGPR.
688 2015.7292621.
- 689 Wright, D., Baltazart, V., Elsworth, N., Hamrouche, R., Karup, J., Lurdes Antunes, M.,
690 McRobbie, S., Mereco, V., Saarenketo, T., 2014. D4.3 monitoring structural and
691 surface conditions (technical report) Eur. Comm. DG Res., Tech. Rep FP7-285119.

692 Zaki, A., Megat Johari, M.A., Wan Hussin, W.M.A., Jusman, Y., 2018. Experimental
693 assessment of rebar corrosion in concrete slab using ground penetrating radar (gpr).
694 International Journal of Corrosion , 1–10doi:10.1155/2018/5389829.