https://doi.org/10.1016/j.conbuildmat.2023.132855

Classification of concrete corrosion states by GPR with machine learning

Phoebe Tin-wai Wong*, Wallace Wai-lok Lai

Department of Land Surveying and Geo-informatics, Faculty of Construction and Environment, Hong Kong Polytechnic University, Hong Kong

Chi-sun Poon

Department of Civil and Environmental Engineering, Faculty of Construction and Environment, Hong Kong Polytechnic University, Hong Kong

Abstract

The evaluation of rebar corrosion in reinforced concrete by using ground penetrating radar (GPR) and machine learning (ML) is a complex process. In this paper, a multi-variate method is presented. It uses full-volume data obtained from the amplitude domain in a regular GPR x-y scanning exercise, and the shape of the rebar's reflection to categorise different corrosion phases. This method allows multi-dimensional analysis with quantifiable GPR attributes. GPR data were extracted from the field and laboratory and then labelled according to the ground truths and reference specimens. A classic ML algorithm, logistic regression, was applied. The cross-validation accuracy (sensitivity and specificity) of individual corrosion phases was high ($> 99\%$), and the false alarm rate was low ($< 1\%$). This work shows that GPR as an evaluation tool can assess unseen data like doing blind tests. Nonetheless, continuous expansion of the training database is suggested to increase its diversity in the future.

Keywords: ground penetrating radar; machine learning; concrete corrosion;

*Corresponding author

Email addresses: tin-wai-phoebe.wong@connect.polyu.hk (Phoebe Tin-wai Wong), wallace.wai.lok.lai@polyu.edu.hk (Wallace Wai-lok Lai), chi-cun.poon@polyu.edu.hk (Chi-sun Poon)

logistic regression

1. Introduction

Steel-reinforced concrete is one of the most cost-effective and widely used materials for the construction of buildings and infrastructure [1]. However, this material suffers from corrosion of the reinforcing steel bars, which in later stages of development can lead to deterioration of the structure, causing defects such as longitudinal cracking, delamination, and spalling. Therefore, the detection of corrosion in reinforced concrete (RC) structures has been an active field of research in structural health monitoring [1, 2, 3, 4].

Previous research described in [5, 6, 7, 8] demonstrated how the presence of moisture and chloride ions change the electromagnetic (EM) properties of the concrete and, as a result, affect the reflected ground penetrating radar (GPR) signals. Authors in [9, 10, 11, 12] assess concrete bridge specimens in the field using GPR. The GPR results of different corrosion phases show distinct characteristics, which shows the potential of using GPR attributes for machine learning (ML). Currently, the application of ML for predicting the health of concrete components based on GPR data usually relies on an image-based analysis of the reflections from the targets [11, 13, 14, 15, 16]. These targets include rebars and pipes in which hyperbolic reflections are returned as (vertical) cross-section images (a.k.a. B-scans). The review papers [15, 17] suggest an unrealised potential in exploring horizontal cross-section images (a.k.a. C-scans) as well, which contain more information in the spatial domain. Given the research problems presented above, this study attempts to develop SML models for classifying the three corrosion states in concrete through a series of experiments and algorithm development. The aim of this contribution is to establish an ML-based workflow for chloride-induced corrosion evaluation by using GPR signal attributes

compiled from C-scans. The goal of the presented research is to adapt machine learning methods to a novel non-destructive testing (NDT) dataset in order to improve the evaluation of the corrosion state of concrete structural components in civil engineering. This paper primarily focuses on marine structures because of their severe exposure condition. It shortens the service life and produces distinct differences between the corrosion phases. This is desirable for taking snapshots of the structure for identifying defects.

In short, this study mimics the medical diagnostic logic, i.e., image, and diagnosis before extracting samples of human bodies, and demonstrates its practicality to the conventional construction industry. This paper demonstrates how to prepare the GPR data and train a supervised classifier with selected corrosion-indicative GPR attributes. The study is also multi-disciplinary in nature as it demonstrates the effectiveness of using a combination of GPR, as a non-invasive geophysical technology, and machine learning techniques to diagnose a major civil engineering problem: chloride-induced corrosion in concrete. It carries significant impacts to ignite the potential of not only GPR but other non-destructive methods for concrete diagnosis and the value of interdisciplinary collaboration in tackling real-world challenges.

2. Literature Review

2.1. Corrosion mechanism measured by GPR

The condition of a concrete component can be classified based on a few corrosion phases. Hong, et al. [6] proposed the following classification (Figure 1 in [6]). The three phases are "intact", "corrosion initiation/chloride intrusion" and "corrosion propagation/active corrosion". There is free expansion caused by the rust particles, which eventually leads to cracking in the final phase. The effects of degradation modes of the GPR signal are summarised in Table 1.

	Amplitude		**Centre Frequency of the Wavelet**	
	Rebar's reflected signal	Direct wave	Rebar's reflected signal	Direct wave
Chloride intrusion	↓	↓	↓	↓
Rust expansion	↑	↑	no change	no change
Surface cracks	↑ (hairline); ↓ (minor/wider)	↑	↑	↑
Delamination	↑	N/A	N/A	N/A

↑: increase, ↓: decrease

Table 1: Effect of different degradation modes on the amplitude and frequency attributes of GPR signals with reference to [5, 6, 7, 8].

The presence of chloride influences both the amplitude and the frequency attributes. When the concrete is intrusion by chlorine solution, the dielectric permittivity (ϵ) and electrical conductivity (σ) increase, thus leading to a returned GPR signal with reduced amplitude and increased time of flight.

The early stage of active corrosion before major rust formation and expansion is difficult to detect with GPR because the reflections from the concrete-steel interface are always more pronounced than the less obvious corrosion products of active corrosion. Nonetheless, the amplitude in the active corrosion phase both increases when there is rust expansion and increases further when internal cracks or delamination are formed. This has been explained in [8] by the increasing entrapped air content in the subsurface yields more total back-scattering energy for the same size of radar footprint. The authors presented experimental results showing that cracks from an opening of more than two mm can lead to a significant reduction in the rebar reflection amplitude.

2.2. From amplitude thresholding to machine learning

Using thresholds for evaluating concrete conditions is a common approach, such as the measurement of half-cell potential (HCP). Several researchers have attempted to establish threshold values relating to rebar reflection amplitude

in association with corrosion in concrete. Barnes, et al. [18] improved an existing amplitude threshold relative to the top 10% maximum amplitude values of the bridge deck by correcting the effect of depth. Sun, et al. [19] proposed an amplitude threshold for assessing areas that begin to deteriorate, which is determined by the median absolute deviation as an outlier detector. Mohammed Abdelkader, et al. [20] used unsupervised ML algorithms to determine the optimal thresholds for classifying "good", "medium", "severe", and "very severe". Wong et al. [8] proposed 3 arbitrary thresholds by referring the rebar reflection amplitudes to ground truths to distinguish "intact", "active corrosion" and "deterioration". Steel corrosion in concrete is an electrochemical process that takes place at the smallest (atomic) level. Its implicit observation by single, relatively coarse signal features is accordingly prone to error. ML and data analysis techniques in general offer the possibility of reducing error by modelling relationships between multiple data resources using empirical data. Conditional and marginal correlations are exploited, which can significantly reduce error compared to uni-variate thresholding.

Earlier studies mainly focused on classifying defects by analysing the rebar reflection amplitude [5, 6, 7, 8] or the shape of hyperbolic reflections of the signals [11, 13, 14, 15, 16] using image analysis of B-scans, while very few applied classic ML in corrosion evaluation [21]. The major limitation with relying on the identification of hyperbolic reflections from rebars is that such a feature is not always present. Studies [22, 8] have shown that delamination between rebars returns deformed, relatively flat, and strong reflections in B-scans.

Travassos, et al. [15] reviewed the application of an artificial neural network (ANN) and ML for classification and decision-making with GPR. A few challenges were highlighted in their conclusion, including the transition of analysis from 2D to 3D for a full volume of information in the spatial domain,

analysis of scenarios with increased complexity, and data fusion with different sensors. Völker, et al.[21] investigated the effectiveness of using multi-variate labelled data collected by four different NDT instruments (including the HCP and GPR) to train a logistic regression (LR) classifier for concrete assessment. GPR A-scan and C-scan information were used to train the classifier. The results showed that the combination of using attributes from HCP and GPR gives the highest accuracy (true positive rate: 0.58, false positive rate: 0). This paper shows that there are other corrosion-indicative GPR attributes instead of only the rebar reflection amplitude. Moreover, LR is easy to implement, train and update. It also performs satisfactorily on data with linearly separable distributions. Nonetheless, the model complexity of an LR classifier can be increased by using polynomial features if desired.

The support vector machine (SVM) is another classic supervised machine learning algorithm. However, SVM is not selected in this paper because LR provides interpretable model coefficients that reveal the importance of each feature. Moreover, LR is not computationally expensive, and the dataset assembled by various GPR attributes is not high-dimensional. Therefore, although SVM is also a powerful supervised algorithm, LR is a more suitable choice for the task and dataset in hand.

Based on the above, this paper attempts to identify various corrosion-indicative GPR attributes for training an LR classifier with polynomial features. The trained classifier is expected to be capable of accurately predicting the respective corrosion phases across the concrete under testing. With the success of the proposed method, GPR can become a standalone evaluation tool for evaluating chloride-induced corrosion.

3. Materials and Data Acquisition

The effects of chloride-induced corrosion and cracks on reflected GPR signals in different concrete scenarios (intact, chloride-intrusion, delaminated, etc.) are investigated through laboratory experiments and field experiments. For this study, data from the laboratory specimens and field structures were compiled (Table 2).

Specimen D2, CEE_Wall_V, F4, Korr1, and Korr_2013 (Mala/GSSI) are laboratory specimens (> 90 days old), while HKM and NP are field structures that have been exposed in the marine environment for more than 30 years at the time of data acquisition. TM is another field structure that is situated in a > 40-year-old industrial building.

Source	Specimen	Time window (ns)	Step size (m)	Concrete cover (mm) (cover-to-top)	Grid spacing (m)	Rebar size (mm)
Laboratory	D2	3	0.003	40, 48	0.1 x 0.1	12
	CEE_Wall_V	10	0.0025	50	0.1 x 0.1	20
	F4	3	0.003	40	0.1 x 0.1	12
	Korr1	10	0.005	45	0.1 x 0.1	12
	Korr_2013 (Mala)	10	0.0017	75	N/A	25
	Korr_2013 (GSSI)	10	0.0025	75	N/A	25
Field	HKM	10	0.003	50	0.1 x 0.1	N/A
	NP	10	0.003	50	0.1 x 0.1	N/A
	TM	6	0.0025	50	0.1 x 0.1	N/A

Table 2: Survey parameter settings for each experiment.

Among the 7 specimens, Korr1 and CEE_Wall_V were measured by GSSI SIR-20 with a 2 GHz antenna. Korr_2013 (Mala) and Korr_2013 (GSSI) were measured by Mala ProEx with a 1.6 GHz antenna and GSSI SIR-20 with a 2.6 GHz antenna in Germany, respectively, from the same concrete. The rest were surveyed by GSSI SIR-4000 with a 2 GHz antenna in Hong Kong. In total, 4 control units and 4 antennas were employed in two different laboratories. A method for harmonising the data sets will be described later.

Specimen D2, F4, Korr_2013 (Mala/GSSI) and Korr1 were all concrete slabs prepared with a typical concrete mix, the water-to-cement ratio ranges from

0.49 to 0.6. Specimen D2 has an artificial wedge-shaped delamination between rebars in the designated area. Rebar samples influenced by the delamination are not used in this study. The specimen was immersed in a 4% saline solution after curing for 90 days in the laboratory. The immersion lasted for 14 days under ambient conditions to imitate surface chloride intrusion. After that, the specimen was lifted and only half of it was in contact with the 4% saline solution, and it was eventually air-dried. Data were captured in between all these states (adding up to 4 total measurement sets.). The size of the measurement field was 0.27 m x 0.36 m. Specimen F4 was immersed in 4% saline for 7 days and then let dry for 14 days. Direct current (DC) was impressed on rebars to accelerate electrochemical corrosion. Surface cracks were produced due to the accumulation of corrosion products. The size of the measurement field was 0.52 m x 0.36 m. Data from chloride intrusion up to the appearance of hairline crack were acquired. The datasets Korr_2013 (Mala) and Korr_2013 (GSSI) come from the same specimen, its details have been reported in [5]. Multi-frequency concrete data (28 hyperbolas in one B-scan) before exposure to chloride were acquired when the specimen was 104 days old. The details of Korr1 have been reported in [21]. The baseline data of this specimen before chloride intrusion are used in this study.

The concrete wall (CEE_Wall) is a rather old laboratory specimen (1.8 m x 2 m) when compared to other laboratory specimens. It was made in 2010. This specimen has always been stored in an indoor and ambient environment. Part of the specimen was mixed with 3% saline to imitate chloride intrusion. The rest of the specimen was prepared by a typical concrete mix (water-to-cement ratio: 0.5). The specimen has 2 sides, there are 5 vertical rebars on Side V, and the upper 50 cm are within the chloride-intruded area. Side V was measured by GSSI SIR-20 in 2012. Only the intact part of Side V is used in this paper.

Specimen NP is a concrete column (1.1 m x 2.7 m) located at a local pier (Figure 1). Specimen HKM is a concrete slab (9 m x 3.9 m) located at a local ferry terminal (Figure 2) in Hong Kong. As both marine structures were built more than 30 years ago and exposed in the splash zone [8, 23], it is reasonable to assume that the concrete was highly chloride-intrusion.

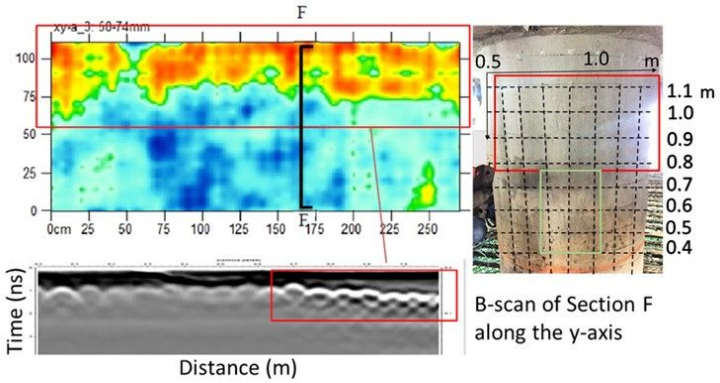

Figure 1: Site photo of NP and selected GPR results.

Concrete cores were extracted from the slab at HKM (Figure 2) after GPR measurement. Delamination was confirmed in the area facing the sea (core C1) and internal cracks (core C3) were frequently observed in the cores. The GPR survey results have been reported in [8, 23] and so the details are not reported here. The size of the measurement field was 3.9 m x 9 m (Figure 2). The mesh grid was 10 cm x 10 cm.

The condition of the NP column was assessed by hammer tapping and visual inspection. It was found that the upper part of the column (y: 0.8 to 1.1 m) has delaminated, and minor surface cracks extended from the delaminated area (1-2 mm) were observed (Figure 1).

A 2.7 m x 3.5 m concrete wall was measured in Tuen Mun (TM). The mesh grid was 10 cm x 10 cm. Cubicle toilets are installed on the other side of the measured wall, and the space on the side being measured was divided into two

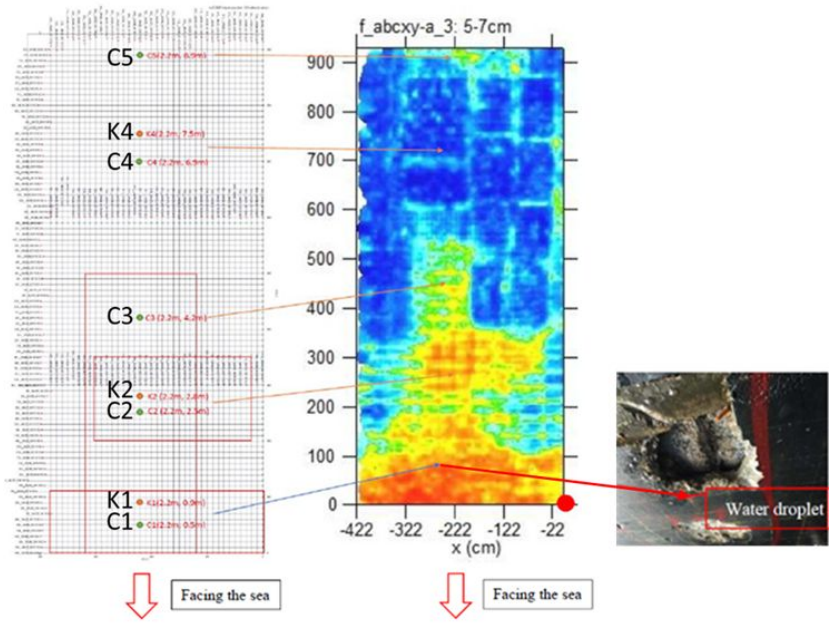

Figure 2: Coring positions for the HKM data and the associated locations on the C-scan [8].

decks due to the high headroom height. Therefore, producing two sets of data acquired separately. Surface cracks and spalling were observed occasionally, and delamination was detected at the bottom of the lower deck by hammer tapping (Figure 3). A series of destructive tests were performed to assess the chloride content, moisture content, and compressive strength. These ground-truths aid the labelling processing of the data sets for training.

Figure 3: The upper deck (left) and the lower deck (right) of the Tuen Mun wall. The red boxes indicate the surface defects and the blue crosses indicate locations with debond.

4. Post-processing and Attribute Extraction

Conventional GPR post-processing procedures were performed before attribute extraction for machine learning. Only the essential procedures (DC shift removal and Butterworth filtering) were kept so that the reflection amplitudes were not modified. Signal gain amplification was removed to assure, that the rebar reflection amplitude across all data sets is compatible for comparison after amplitude conversion. Migration of the measured data was not conducted in this study because the amplitude of the direct wave and reflected wave stored within the dynamic range would be lost.

Next is the generation of C-scans. The locations of the direct waves (DW) and reflection waves (RW) in the time domain (or radargrams) were identified, and their amplitudes were then extracted and bi-linearly interpolated to produce amplitude slices of the DW and RW. Out of the many attributes that have been studied before (Table 1), such as the velocity variation due to diffraction hyperbola, amplitude attributes are found to be the most appropriate. Frequency attributes were not chosen because literature (Table 1) shows that the frequency attributes tend to be less sensitive in the characterisation of all 3 phases in comparison to the DWA and RWA. Therefore, these two attributes are important corrosion indicators.

Figure 4 summarises the workflow for preparing the data and attribute extraction. As shown in the flowchart, two GPR features, i.e. the strong flat reflection layer and the hyperbolic reflection from rebars, were identified from the radargrams after 2D processing. The former could be delamination, back-wall reflection, or from the background. The latter represents rebars. Figure 5 provides two examples of flat reflection layers associated with corrosion in concrete. In contrast, deterioration is not likely if the hyperbolic reflection is clearly identifiable. Hence, the presence of a strong flat layer is an important

corrosion indicator. Together with the two amplitude attributes, there is a total of 3 corrosion indicators for supervised machine learning (SML) in this paper (Table 3).

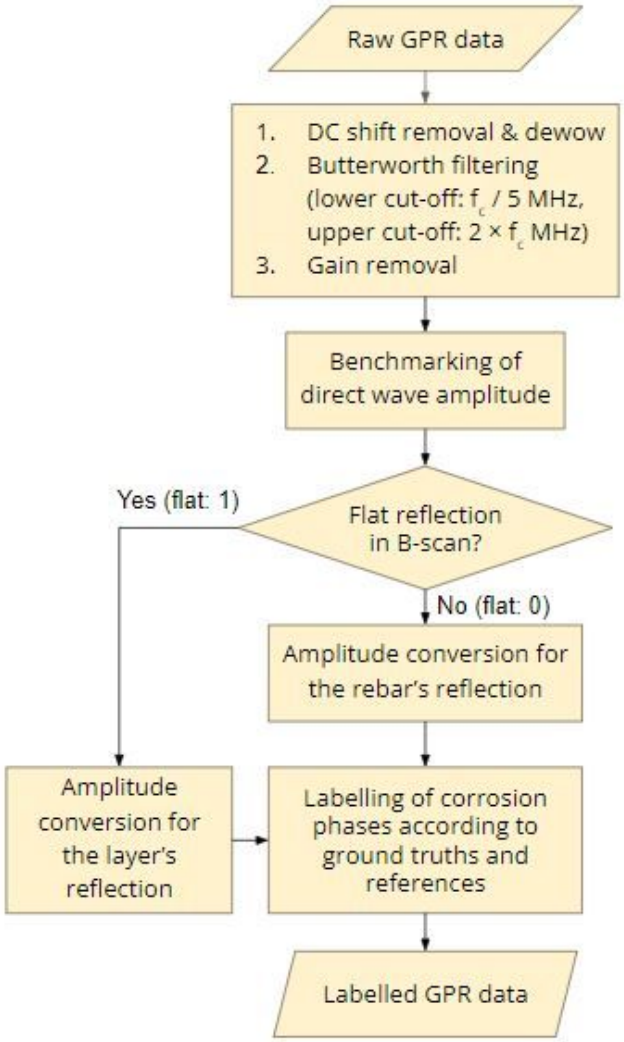

Figure 4: Workflow for GPR data post-processing and attribute extraction.

The following two steps are designed. First, for locations with rebar, the amplitudes were extracted from the RW of the rebar and the associated DW, and the "flat" variable is "0" since the hyperbolic reflection is clear. Second, for

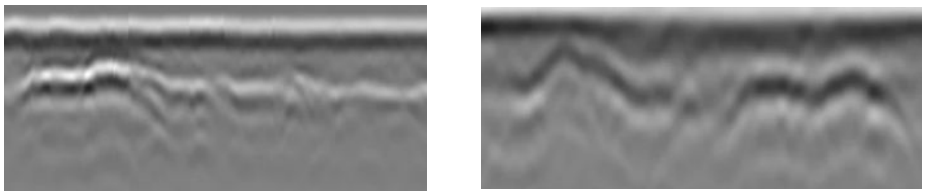

Figure 5: Examples of strong flat reflection layer from delamination (left) and delamination with out-of-phase wavelets (right) in concrete.

Corrosion Indicator/ Attribute	Type	Description	Unit
1	Image-based	Strong flat reflection layer observed in B-scan (0: not seen, 1: seen)	-
2	Ray-based	Reflection amplitude of rebar/layer (RWA)	dB
3		Direct wave amplitude (DWA)	dB

Table 3: List of the attributes that were extracted from the GPR data

locations with a strong flat layer and no rebar, the amplitudes were extracted from the RW of the layer and the associated DW, the "flat" variable is "1" since the A-scan is associated with the strong flat layer. If the polarity of the RW is not positive, the polarity of the RW was reversed before attribute extraction. This is because the positive part of the wavelet was extracted during the picking of amplitude values from A-scans. The polarity of the direct wave and reflected wave is instrumentation-specific, therefore, the polarity might need to be reversed in order to correctly extract the amplitude information.

Harmonisation of the amplitudes acquired from different GPR systems is required, such that the amplitude attributes are compatible for assembling into a single fingerprinting database for machine learning. A reference GPR instrumentation and reference concrete specimen were used to provide two universal and empirically calibrated constants (k_1 and k_2) for modifying the conversion equation (Eq. 1) described in ASTM D6087 [24]. The details of the development of this harmonisation method are described in [25]. In Eq. 2, k_1 is for deriving the scaling factor. It comes from the mean DWA of a reference specimen, an

intact and dry concrete, acquired with the reference GPR instrumentation. Intact specimens are suitable for benchmarking the converted amplitudes of the returned signals because their EM properties are relatively homogeneous. The scaling factor (m) bridges different datasets obtained by the instrumentation that needs to be harmonised by referring to the mean DWA of the concrete under testing. k_2 is the average amplitude of the direct wave of the reference antenna probe in air measured by the reference GPR instrumentation. The cover depth (D) was estimated based on the estimated velocity [26].

$$\text{Amplitude Conversion (to dB): } A_{dB} = 20 \cdot log_{10}\left(A_i\right) \tag{1}$$

Eq. 1 was modified to take the attenuation effect due to depth variation of rebar and the scaling between different GPR systems into account, Eq. 1 becomes:

$$A_{dB} = 20 \cdot log_{10}\left(A_R \cdot e^D \cdot m/k_2\right) \tag{2}$$

where A_i is the measured amplitude within the GPR dynamic range, A_0 is the direct wave amplitude, A_R is the reflection amplitude from rebars or strong flat layers, $m = k_1/\bar{A_{DW}}$ is the scaling factor for conversion, k_1 and k_2 are universal and empirical calibrated constants, $\bar{A_{DW}}$ is the mean DWA of the reference concrete measured by the instrumentation that needs to be harmonised, $D > 0$ is the cover depth to rebars or strong flat layers (m). When $D = 0$, which $e^D = 1$ becomes a special case representing the DWA.

After amplitude conversion, the three GPR attributes are assembled and annotated with reference to the ground truths and laboratory references in Tab. 2. Since corrosion is a continuous process, it is categorised into 3 phases for annotation and easy understanding of measures to be taken (Table 4).

Phase	Civil Engineering Example
1. Normal	Functioning normally; the structural integrity is as expected according to the design service life model
2. Depassivation or commencement of corrosion	Depassivation induced by chloride intrusion; rust expansion/early corrosion
3. Deterioration	Defects have developed; internal cracks or delamination or surface cracks are present

Table 4: Description of each corrosion phase. The phases are indexed into 1, 2 and 3 for simplification.

The data selected for labelling and training are listed in Table 5. Specimen D2 provides intact and chloride-intrusion concrete data (D2_0 and D2_1). Another laboratory specimen F4 provides data on depassivation and rust expansion (F4_2). As described earlier, the datasets Korr1-0, Korr_2013 (Mala), Korr_2013 (GSSI), and CEE_V_0 were measured by 16-bit systems (GSSI SIR-20 and Mala ProEx). The data extracted from the intact area were used for training the classifier.

Source	Dateset	Phase 1	Phase 2	Phase 3	Scenario
		functioning normally; intact	chloride intrusion; commencement of corrosion	deterioration	
Laboratory at HK PolyU	D2_0	✓			normal
	D2_1		✓		chloride
	CEE_V_0	✓			normal
	Korr1_0	✓			normal
	F4_2		✓		depassivation; rust expansion
Laboratory at BAM	Korr_2013 (Mala)	✓			normal
	Korr_2013 (GSSI)	✓			normal
HK-Macau Ferry Terminal	HKM_C1			✓	delamination
	HKM_C3			✓	internal crack
North Point Vehicular Pier	NP_D			✓	delamination
	NP_SC			✓	minor surface cracks
Tuen Mun (on the other side of cubicle toilets)	TM1_C1		✓		water seepage
	TM2_C5	✓			normal
	TM2_C6			✓	corroded
	TM2_C8			✓	corroded; section loss
	No. of A-scans	7586	3786	5411	

Table 5: Summary of training data selected for machine learning.

The remaining training data come from field structures. 4 sets (HKM_C1,

HKM_C3, NP_D, NP_SC) of size 10 cm x 10 cm were selected from the HKM and NP data sets for training. Real defects in marine structures are provided by these data sets and ground truths or references are available for labelling. The last 4 sets of field data, each of size 10 cm x 10 cm, come from an internal wall on the other side of cubicle toilets. Core C1 from TM1 is confirmed to be suffering from severe water seepage and is therefore labelled as Phase 2. TM2-C5 is considered intact based on the destructive test results. TM2-C6 is labelled as Phase 3 because it is right below the spalled and crack area, and the internal structure is too weak for mounting the coring structure. TM2-C8 is also labelled as Phase 3 because the rebar has disappeared upon core drilling. Data of size 10 cm x 10 cm were extracted from these 4 areas, respectively.

5. Establishment of Machine Learning Classifier

Figure 6 describes the workflow for establishing the SML classifier with the fingerprinting database. First of all, feature rescaling of the continuous variables, which are the amplitude attributes, was carried out by using the robust scaler (Eq. 3) [27]. The robust scaler makes use of the inter-quartile range (Q_3 - Q_1) to rescale the data (Eq. 3) statistically, therefore, the effect of outliers is minimised.

$$x_{scaled} = \frac{x_i - Q_1(x)}{Q_3(x) - Q_1(x)} \tag{3}$$

where x_{scaled} is the rescaled DWA and RWA data, x_{train} is the data from the training set, x_{test} is the data from the test set, SD is the standard deviation, Q_3 is the 75^{th} quantile, and Q_1 is the 25^{th} quantile.

The rescaled database was then used to train the logistic regression classifier (Eq. 4 - 5), the parameters (β) were optimised by determining the gradient descent, the regularisation strength (C) and the degree of model complexity (d)

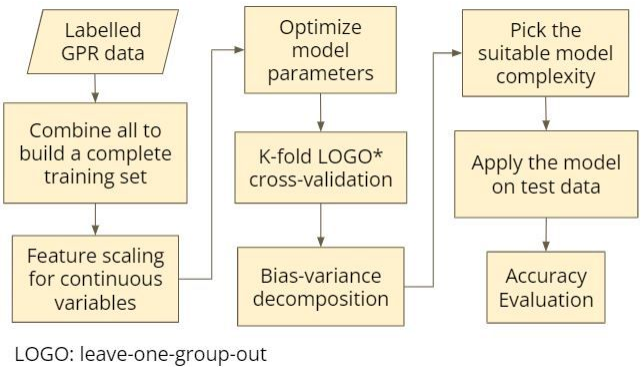

Figure 6: Workflow for supervised classification on labelled GPR data.

in order to generalise the decision boundaries for the three corrosion phases.

$$y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_n X_n \tag{4}$$

$$P = 1/\left(1 + e^{-y}\right) \tag{5}$$

where y is the hypothesis function, β is the parameter of the hypothesis, and P is the probability of success.

Optimisation of the classifier was aided by bias-variance decomposition and k-fold cross-validation. A range of regularisation strengths (C: 0.1, 0.2, 0.4, 0.6, 0.8, 1, 2), and a range of degree of polynomials (d: 1 to 7) were tried for the LR classifier. The "L2" penalty was used in regularisation. The errors of the optimised classifier itself were assessed by bias-variance decomposition [28]. This step was integrated with leave-one-group-out (LOGO) k-fold cross-validation (Figure 7). Figure 7 illustrates how cross-validation was executed for this study by splitting and re-assembling each data set in Table 5.

The group with the best fit was selected ultimately for training. The class weight (Eq. 6) [23] was included to address the issue of imbalanced datasets between different phases.

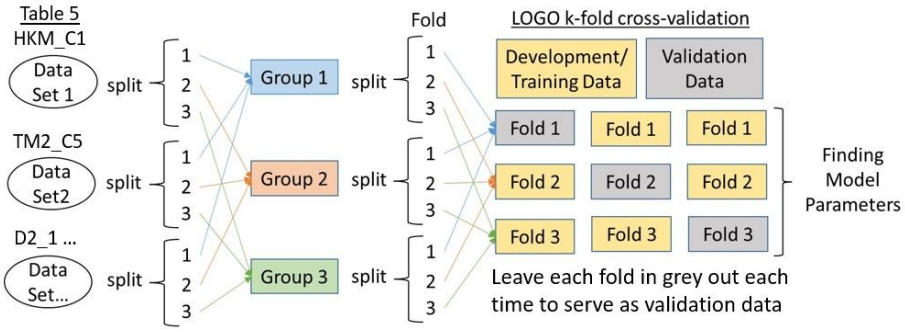

Figure 7: Description of LOGO k-fold cross-validation for this study.

$$weight = \frac{n_{sample}}{n_{class} \times n_{y,\ occurrence}} \tag{6}$$

where n_{sample} is the number of samples, $n_{sample}$is the number of classes, and $n_{y,occurrence}$ is the count number of occurrences of each value of the respective label/class (i.e., y).

The accuracy of the classifier is represented by the class-weighted f1-score, which is expressed as the development score and validation score in bias-variance decomposition [28], respectively. For individual corrosion phases, the sensitivity (true positive rate), the specificity (true negative rate), and the false alarm rate (false positive rate) are presented.

6. Findings and Discussion

6.1. Training Dataset

Figure 8 shows the three selected GPR attributes in the feature space (Table 3) as 2D cross-plots to represent the 3D feature space. A "tail" is observed in the DWA-RWA plot (Plot 1 or 3 in Figure 8), that part of the data instances belongs to HKM-C3, which represents the area with a very fine internal crack inside the concrete slab at the pier. Another "tail" is the TM-C1, in which the RWA is significantly low due to water seepage. In addition, the green cluster that

overlaps with the Phase 1 (or blue) cluster comes from TM2-C8, the cluster below it is TM2-C6. They represent corroded concrete with lower DWA. In addition, the attribute "flat" can effectively separate Phase 3 data associated with the flat reflection layer from the overlapping Phase 1 data. It is clear that neither thresholding nor linear decision boundaries are capable of distinguishing the 3 corrosion phases accurately (Table 4), and machine learning can compute the optimal decision boundaries for classification.

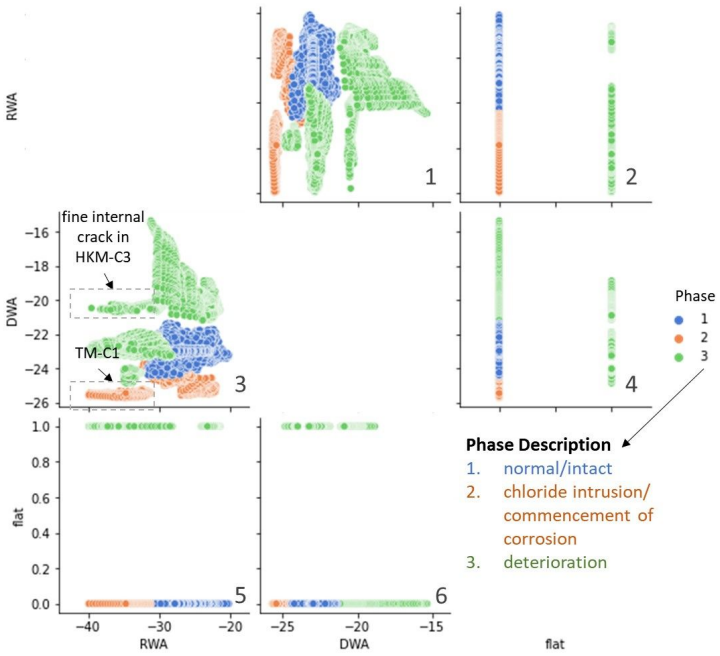

Figure 8: Labelled training data plotted in the feature space.

6.2. Cross-validation

The bias-variance decomposition of Groups 2 and 3 (described in Figure 7) failed to return results with a good-fit curve, therefore, only Group 1 (No. of A-scans: 5,411) was selected for further establishment (data available at [29]). Its decomposition result is displayed in Figure 9 (left). We can see that the development score of Group 1 becomes stable and the validation score has a

decreasing trend after reaching the third order. It shows that the model is technically a good fit. The regularisation strength (C) was chosen based on the bias-variance decomposition with a range of C (described in Section 5) when the order of the polynomial (d) is 3 (Figure 10). Therefore, C=0.8 was ultimately selected.

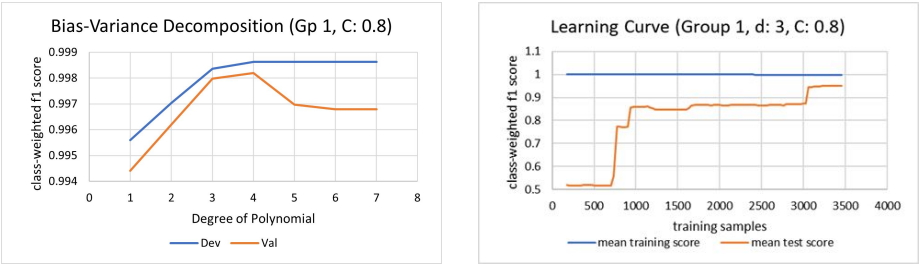

Figure 9: (left) Bias-variance decomposition of the LR classifiers of Group 1 when C=0.8 and (right) the learning curve of Group 1.

The learning curves of the classifier (Figure 9, right) show that the model is slightly overfitting because the training score decreases very narrowly. For the validation score, it requires around 700 samples to reach a score of 0.8, and more than 3,000 samples to reach a score of 0.9.

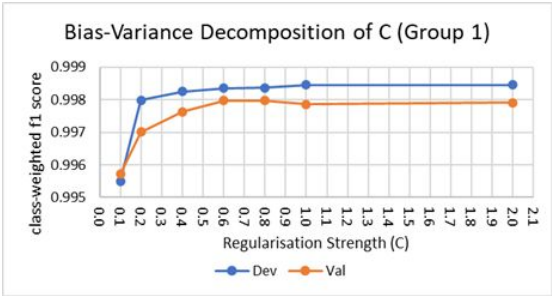

Figure 10: Bias-variance decomposition of the regularisation strength (C) when the degree of polynomial=3.

Figure 11 shows the decision boundary (DB) of each corrosion phase. The blue region belongs to Phase 1, the orange region belongs to Phase 2, and the green region belongs to Phase 3. The definition of each phase has been described

in Table 4. The coefficients or model parameters (Figure 12) of the DBs indicate the importance of each variable. The values of model coefficients are tabulated in Appendix A. From Figure 11 and Figure 12, the two amplitude attributes (RWA and DWA) are shown to be more influential on decision making, and DWA is slightly more dominant than RWA. The scaling factor for amplitude conversion is determined by referring to the DWA of intact concrete, as chloride intrudes, the surface condition of the concrete is no longer homogeneous, and the RWA will also be affected. Moreover, the increase in subsurface air content usually affects the RWA more significantly. In short, both the literature and the experiments in this study show that the effect of corrosion affects the amplitude attributes.

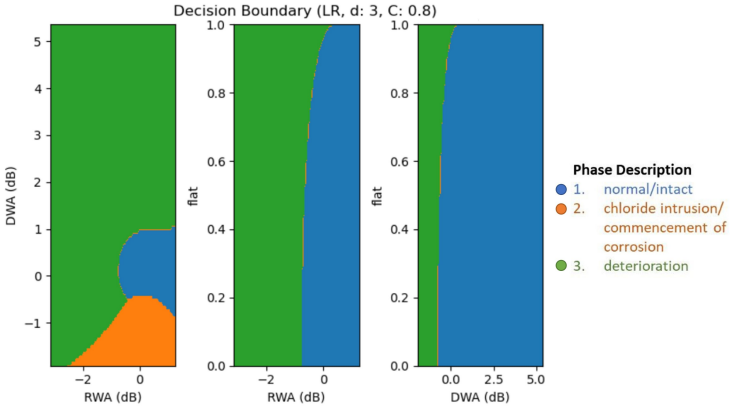

Figure 11: Decision boundary of Group 1.

The decision boundary of Phase 2 in Plot 1/3 (chloride intrusion and commencement of corrosion) concentrates in the lower region, and the decision of Phase 1 (normal or intact) is located at the higher DWA and RWA region. This aligns with the literature and the results of the experiments. The distribution of Phase 3 (deterioration) data does not have a specific pattern, this is because the subsurface of deteriorated areas is complex, and the returned signals could have high DWA and low RWA (e.g., the wall in Tuen Mun), or high DWA and

RWA (e.g., delamination at HKM). As a result, the decision boundary of Phase 3 occupies such a large area in Plot 1/3. In addition, Figure 12 (bottom) reveals the importance of the "flat" variable. This variable is among the top 5 model coefficients in the decision boundary of Phase 3. Since flat=1 only exists in Phase 3, this implies that the "flat" variable helps separate the overlapping amplitudes between Phase 1 and 3. To sum up, the decision boundary of each phase exhibits a similar pattern that can be observed in reality. This is important because a classification model that can represent the physical world is more capable of dealing with unseen test data.

Table 6 details the evaluation metrics of individual corrosion phases (described in Table 4) from cross-validation. We can see that the overall cross-validation accuracy of the classifier is high. This is because both the sensitivity and specificity of the 3 phases are high, and their false alarm rates are very low. The FPR of Phase 3 being 0 implied that the falsely classified samples of Phases 1 and 2 are not Phase 3. It means that the deteriorated areas (Phase 3) are not likely overlooked.

Phase	1 (normal or intact)	2 (chloride intrusion or commencement of corrosion)	3 (deterioration)
Sensitivity (%)	99.684	99.724	100
Specificity (%)	99.896	99.815	100
False Alarm Rate (%)	0.104	0.185	0

Table 6: Evaluation metrics of individual phases of the training database.

7. Conclusion

To conclude, GPR does not only provide an imaging technique, but it also generates useful and quantifiable information for defect classification and diagnosis supported by ground-truthing or pre-determined stages in laboratory

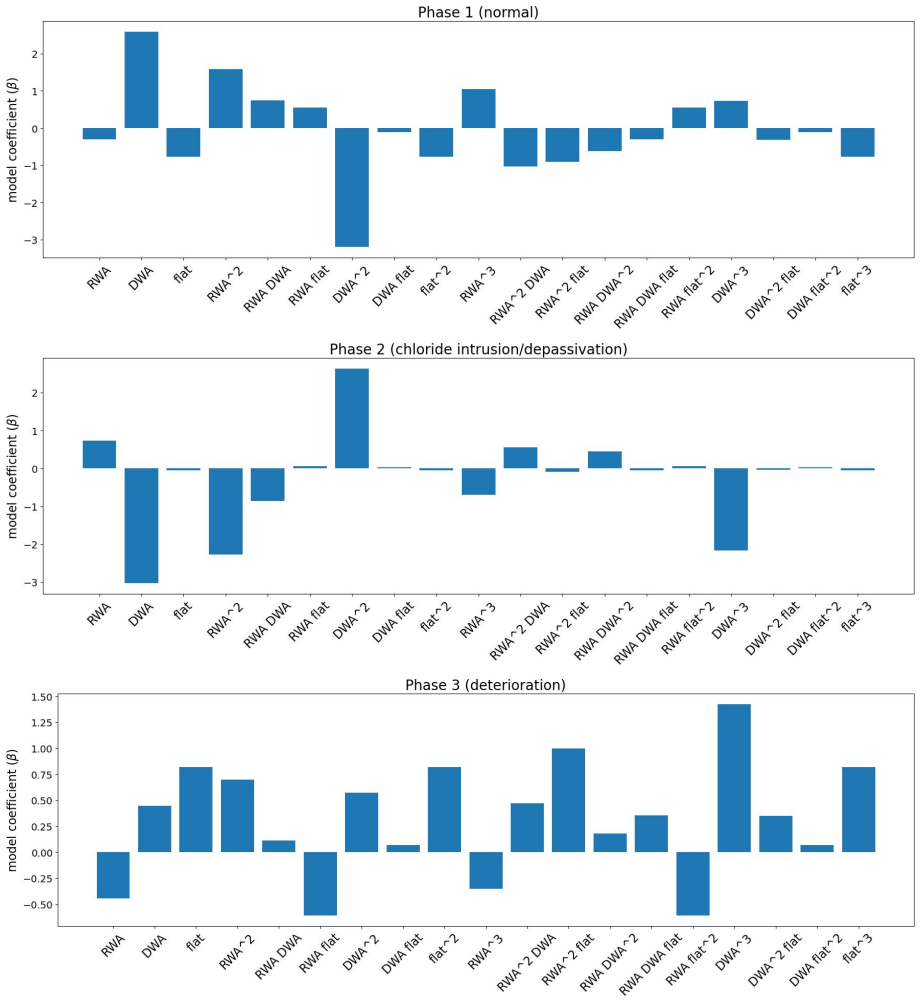

Figure 12: Coefficients of the decision boundary function of Group 1 in Eq. 4 .

experiments. A post-processing workflow for extracting the GPR attributes is developed for producing compatible GPR attributes for establishing a "fingerprinting" database. The 3 clusters of data instances (all: 16,783 and Group 1: 5,411 A-scans) of the corrosion phases are moderately separated in the feature space.

By following the suggested GPR-ML workflow, an optimised SML classifier

making use of the LR algorithm was applied. Cross-validation shows that the performance of the LR classifier works well on seen data. The results in the paper demonstrate the potential applicability of the proposed workflow. The fingerprinting database expressed in the feature space exhibit the limitation with uni-variate thresholding, where the data instances of the 3 corrosion phases largely overlap. The performance of the established classifier on unseen test data will be examined in another article. Meanwhile, continuous expansion of the fingerprinting database with concrete data from different scenarios will also be carried out for enhancing the robustness of the classification model. Further investigations are also required to evaluate whether the model is ready for field application.

8. Acknowledgements

The authors would like to thank the following bodies: The Civil Engineering and Development Department (CEDD) of the Hong Kong SAR Government for providing the field sites (HKM, NP), the professional advice on machine learning provided by Dr Christoph Völker of BAM, Germany, the Civil and Environmental Engineering Department of HKPolyU with the preparation of some of the specimens, and the English editing service provided by Dr Mick Atha.

This study was financially supported by the following projects: General Research Fund of the Research Grant Council of Hong Kong, Project PolyU/15216619 'Waveguide Characterization of Shallow Subsurface Damages in Infrastructure Materials by Dispersive Behavior of Ground Penetrating Radar Wave'; Smart Transport Fund of HKSAR Government, Project PSRI/14/2109/RA 'Smart Assessment of Bridge Deck Efficiency and Safety in Hong Kong'

Appendix

A. Model Coefficients of Group 1's Decision Boundaries

Phase	y-intercept	RWA	DWA	flat	RWA^2	RWA·DWA	RWA·flat
1	4.478852	-0.2956	2.585048	-0.77258	1.57771	0.73894	0.546341
2	-1.16569	0.737484	-3.03223	-0.0463	-2.27672	-0.85366	0.063408
3	-3.31317	-0.44188	0.447186	0.818878	0.699011	0.114715	-0.60975

Table A.1: Model coefficients of the decision boundaries (Part 1)

Phase	DWA^2	DWA·flat	$flat^2$	RWA^3	RWA^2·DWA	RWA^2·flat	RWA·DWA^2
1	-3.1985	-0.1033	-0.77258	1.047214	-1.02961	-0.90964	-0.62368
2	2.627987	0.036131	-0.0463	-0.69713	0.558667	-0.08818	0.444024
3	0.570508	0.067166	0.818878	-0.35008	0.470947	0.997826	0.179656

Table A.2: Model coefficients of the decision boundaries (Part 2)

Phase	RWA·DWA·flat	RWA·$flat^2$	DWA^3	DWA^2·flat	DWA·$flat^2$	$flat^3$	RWA·DWA^2
1	-0.30482	0.546341	0.73686	-0.31577	-0.1033	-0.77258	-0.62368
2	-0.05146	0.063408	-2.15929	-0.03251	0.036131	-0.0463	0.444024
3	0.35628	-0.60975	1.422432	0.348283	0.067166	0.818878	0.179656

Table A.3: Model coefficients of the decision boundaries (Part 3)

References

[1] A. M. Neville, Properties of concrete, 4th Edition, Longman Group, Harlow, 1995.

[2] J. P. Broomfield, Corrosion of steel in concrete: understanding, investigation and repair, 1st Edition, CRC Press, London, 1997.

[3] F. Hunkeler, Corrosion in reinforced concrete: processes and mechanisms, CRC Roca Raton, FL, 2005. doi:10.1533/9781845690434.1.

[4] A. Poursaee, Corrosion of steel in concrete structures, Elsevier, Woodhead Publishing, Duxford, UK, 2016.

[5] W. L. W. Lai, T. Kind, M. Stoppel, H. Wiggenhauser, Measurement of accelerated steel corrosion in concrete using ground-penetrating radar and

a modified half-cell potential method, Journal of Infrastructure Systems 19 (2) (2013) 205–220. doi:10.1061/(ASCE)IS.1943-555X.0000083.

[6] S. Hong, W.-L. Lai, R. Helmerich, Experimental monitoring of chloride-induced reinforcement corrosion and chloride contamination in concrete with ground-penetrating radar, Structure and Infrastructure Engineering 11 (1) (2014) 15–26. doi:10.1080/15732479.2013.879321.

[7] S. Hong, W. L. W. Lai, G. Wilsch, R. Helmerich, R. Helmerich, T. Günther, H. Wiggenhauser, Periodic mapping of reinforcement corrosion in intrusive chloride contaminated concrete with gpr, Construction and Building Materials 66 (2014) 671–684. doi:10.1016/j.conbuildmat.2014.06.019.

[8] T. P. Wong, W. L. W. Lai, F. C. J. Sham, C. S. Poon, Hybrid non-destructive evaluation methods for characterizing chloride-induced corrosion in concrete, NDT and E Internationaldoi:10.1016/j.ndteint.2019.05.008.

[9] K. Dinh, N. Gucunski, J. Kim, T. H. Duong, Understanding depth-amplitude effects in assessment of gpr data from concrete bridge decks, NDT & E International 83 (2016) 48–58. doi:10.1016/j.ndteint.2016.06.004.

[10] K. Dinh, N. Gucunski, J. Kim, T. H. Duong, Method for attenuation assessment of gpr data from concrete bridge decks, NDT & E International 92 (2017) 50–58. doi:https://doi.org/10.1016/j.ndteint.2017.07.016. URL http://www.sciencedirect.com/science/article/pii/S0963869517301949

[11] K. Dinh, N. Gucunski, T. H. Duong, An algorithm for automatic localization and detection of rebars from gpr data of concrete bridge decks, Automation in Construction 89 (2018) 292–298.

doi:https://doi.org/10.1016/j.autcon.2018.02.017.
URL http://www.sciencedirect.com/science/article/pii/S0926580517307136

[12] K. Dinh, N. Gucunski, T. Zayed, Automated visualization of concrete bridge deck condition from gpr data, NDT & E International 102 (2019) 120–128. doi:https://doi.org/10.1016/j.ndteint.2018.11.015.
URL http://www.sciencedirect.com/science/article/pii/S0963869518303451

[13] H. Harkat, A. Ruano, M. G. Ruano, S. D. Bennani, Classifier design by a multi-objective genetic algorithm approach for gpr automatic target detection, IFAC-PapersOnLine 51 (10) (2018) 6. doi:10.1016/j.ifacol.2018.06.260.

[14] U. Ozkaya, F. Melgani, M. Belete Bejiga, L. Seyfi, M. Donelli, Gpr b scan image analysis with deep learning methods, Measurement 165 (2020) 107770. doi:https://doi.org/10.1016/j.measurement.2020.107770.
URL http://www.sciencedirect.com/science/article/pii/S0263224120303080

[15] X. L. Travassos, S. L. Avila, N. Ida, Artificial neural networks and machine learning techniques applied to ground penetrating radar: A review, Applied Computing and Informaticsdoi:https://doi.org/10.1016/j.aci.2018.10.001.
URL http://www.sciencedirect.com/science/article/pii/S2210832718302266

[16] N. Smitha, V. Singh, Target detection using supervised machine learning algorithms for gpr data, Sensing and Imaging 21 (1) (2020) 11. doi:10.

1007/s11220-020-0273-8.

URL https://doi.org/10.1007/s11220-020-0273-8

[17] Z. Tong, J. Gao, D. Yuan, Advances of deep learning applications in ground-penetrating radar: A survey, Construction and Building Materials 258 (2020) 120371. doi:https://doi.org/10.1016/j.conbuildmat.2020.120371.

URL http://www.sciencedirect.com/science/article/pii/S095006182032376X

[18] C. L. Barnes, J.-F. Trottier, D. Forgeron, Improved concrete bridge deck evaluation using gpr by accounting for signal depth–amplitude effects, NDT & E International 41 (6) (2008) 427–433. doi:10.1016/j.ndteint.2008.03.005.

[19] H. Sun, S. Pashoutani, J. Zhu, Nondestructive evaluation of concrete bridge decks with automated acoustic scanning system and ground penetrating radar, Sensors (Basel) 18 (6) (2018) 1955. doi:10.3390/s18061955.

[20] E. Mohammed Abdelkader, M. Marzouk, T. Zayed, An optimization-based methodology for the definition of amplitude thresholds of the ground penetrating radar, Soft Computing 23 (22) (2019) 12063–12086. doi:https://doi.org/10.1007/s00500-019-03764-3.

[21] C. Völker, S. Kruschwitz, G. Ebell, A machine learning-based data fusion approach for improved corrosion testing, Surveys in Geophysicsdoi:https://doi.org/10.1007/s10712-019-09558-4.

[22] K. Dinh, N. Gucunski, Factors affecting the detectability of concrete delamination in gpr images, Construction & building materials 274 (2021) 121837. doi:10.1016/j.conbuildmat.2020.121837.

[23] J. F. C. Sham, W. L. L. Wallace, Diagnosis of reinforced concrete structures by ground penetrating radar survey - case study, in: 2017 9th International Workshop on Advanced Ground Penetrating Radar (IWAGPR), 2017, pp. 1–6. doi:10.1109/IWAGPR.2017.7996058.

[24] ASTM, ASTM D6087-08: Standard Test Method for Evaluating Asphalt-Covered Concrete Bridge Decks Using Ground Penetrating Radar, ASTM (2015).

[25] P. T.-W. Wong, W. W. Lai, Amplitude normalization for gpr data acquired from different instrumentations, in: 19th International Conference on Ground Penetrating Radar, 2022, pp. 131–134. doi:10.1190/gpr2022-024.1.

[26] K.-W. P. Lau, W.-Y. B. Cheung, W.-L. W. Lai, F.-C. J. Sham, Characterizing pipe leakage with a combination of gpr wave velocity algorithms, Tunnelling and underground space technology 109. doi:10.1016/j.tust.2020.103740.

[27] Scikit-Learn (2019). [link].
URL https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.RobustScaler.html

[28] G. James, D. Witten, T. Hastie, R. Tibshirani, An Introduction to Statistical Learning: With Applications in R, New York, NY: Springer, New York, NY, 2021.

[29] P. T.-W. Wong, Ground penetrating radar (gpr) corrosion data (wong, 2023), Mendeley Data (2023). doi:10.17632/wbdr5pdxbd.1.
URL http://dx.doi.org/10.17632/wbdr5pdxbd.1