Proceedings of the 7th International Young Geotechnical Engineers Conference – Scott (Ed.)
© 2022 Australian Geomechanics Society, Sydney Australia, ISBN 978-0-9946261-5-8

Adaptive training of convolutional neural networks for slope reliability analysis in spatially variable soils

Apprentissage adaptatif de réseaux de neurones convolutifs pour l'analyse de la fiabilité des pentes dans des sols aux propriétés spatialement variables

Ze Zhou Wang & Siang Huat GOH
Department of Civil & Environmental Engineering, National University of Singapore, Singapore
e0054291@u.nus.edu

Sai G. S. PAI
Future Health Technology, Singapore-ETH Centre, Singapore; formerly with Future Cities Laboratory, Singapore-ETH Centre, Singapore

ABSTRACT: The random field finite element method (RF-FEM) incorporates spatial variability of soil properties with advanced capabilities of the finite element method for analyzing complex geotechnical systems. However, this method requires extensive computational efforts, which can be alleviated by metamodeling techniques. In this paper, a novel deep learning-based technique that involves Convolutional Neural Networks (CNNs) and an adaptive training strategy are proposed to facilitate RF-FEM analyses. Random fields can be considered as "images" that describe the spatial randomness of soil properties. CNNs can predict quantities obtained using RF-FEM analyses by learning features that contain information about the random variabilities in both spatial distribution and intensity of soil characteristics. Therefore, CNNs can be used as metamodels to replace expensive random field finite element models. A multi-layered ***c-ϕ*** slope case study is used to illustrate this technique. The results show that (i) the trained CNNs accurately predict the values of factor of safety and failure probability, and (ii) the adaptive training strategy, compared with the non-adaptive training strategy, improves the performance of the CNNs. In this regard, the use of CNNs and the adaptive training strategy for efficient slope reliability analyses in spatially variable soils is effectively validated.

RÉSUMÉ : La méthode des éléments finis à champ aléatoire (RF-FEM) intègre la variabilité spatiale des propriétés des sols avec les capacités avancées de la méthode des éléments finis pour analyser des systèmes géotechniques complexes. Cependant, cette méthode nécessite des efforts de calcul considérables, qui peuvent être allégés par des techniques de méta-modélisation. Dans cet article, une nouvelle technique d'apprentissage profond plus efficace computationnellement, faisant appel à des réseaux neuronaux convolutifs (CNN), est proposée pour faciliter les analyses RF-FEM. Une stratégie d'apprentissage adaptif est également adoptée dans l'étude afin d'améliorer les performances des CNN. Les champs aléatoires peuvent être considérés comme des "images" qui décrivent le caractère spatialement aléatoire des propriétés du sol. Les CNN peuvent prédire les résultats obtenus par les analyses RF-FEM en apprenant des caractéristiques qui contiennent de l'information sur les variabilités aléatoires de la distribution spatiale et de l'intensité des propriétés du sol. Par conséquent, les CNN pourraient être utilisés comme métamodèles pour remplacer les coûteux modèles d'éléments finis à champ aléatoire. Une étude de cas de pente ***c-ϕ*** multicouche est utilisée pour illustrer la technique proposée. Les résultats montrent que (i) les CNN entraînés donnent des résultats précis, et (ii) la stratégie d'apprentissage adaptatif améliore la précision avec un coût de calcul inférieur à celui de la stratégie d'apprentissage non adaptatif.

KEYWORDS: slope stability; spatial variability; reliability; convolutional neural networks; metamodel.

1 INTRODUCTION

The spatial variability of physical and mechanical properties of natural soils is arguably the most widely recognized source of uncertainty that can impact the performance of geotechnical systems (Griffiths et al. 2009; Goh et al. 2019; Zhang et al. 2021a). As a result, the Random-Field Finite Element Method (RF-FEM), which integrates the random field theory (Vanmarcke 2010) and the advanced capabilities of the finite element method, has become more widespread in the geotechnical engineering community.

Monte-Carlo simulation (MCS) is a commonly used technique for uncertainty propagation. This technique starts by carrying out the geotechnical analyses for X simulated samples. The number of failure samples $X_{failure}$ is then obtained by comparing the sample results against some pre-defined performance functions, following which the value of probability of failure ($P_{failure}$) can be calculated as $X_{failure}/X$. The Coefficient of Variation (CoV) of $P_{failure}$ (Ang & Tang 2007) can also be calculated to assess the uncertainty associated with the calculated value of $P_{failure}$. The method of Monte-Carlo Simulation, although versatile and conceptually simple, suffers from a lack of efficiency because a large number of samples are usually required to ensure that the calculated $P_{failure}$ converges and a reasonably low value of CoV of $P_{failure}$ is attained, especially for cases involving low levels of failure probability.

Accordingly, many strategies, such as Subset Simulation (Au & Beck 2001; Gao et al. 2019), importance sampling (IS) (Liu et al. 2019) and metamodel-based methods (Jiang and Huang 2016; Li et al. 2019; Wang et al. 2020), have also been proposed for an efficient reliability analysis of a geotechnical system.

Among these strategies, the metamodel-based methods have been shown to be efficient for reliability analysis of geotechnical systems. The core concept of these methods is the use of an approximate model, often referred to as a metamodel, a response surface, or a surrogate model, to replace the computationally rigorous, but expensive finite element analysis. However, the presence of spatial variability adds to the complexity in the construction of the metamodel. This results from the significant growth of the dimensionality of the problem. Although the Polynomial Chaos Expansion (PCE) method, a commonly used metamodel for such problems, has been shown to be effective, the computational efficiency can diminish quickly as the size and dimensionality of the geotechnical system under study grows

(Jiang & Huang 2016; Li et al. 2019).

Recently, Wang & Goh (2021) proposed a new metamodel for geotechnical reliability analysis in spatially variable soils. This technique involves the use of Convolutional Neural Networks (CNNs) to automatically extract and learn features pertaining to the random variabilities in both spatial distribution and intensity of soil characteristics. In this way, CNNs interpret random fields in a completely different way from that adopted by the conventional PCE-based techniques. After sufficient training using a set of sample data, CNNs gain the capability to make good quality predictions in lieu of the finite element analysis. Most importantly, Wang et al. (2021) demonstrated that this technique is robust against the growth of the dimensionality of the problem, which makes it a better metamodel than a PCE.

The primary objective of this paper is to investigate use of an adaptive training strategy to improve the performance of CNNs for slope reliability analyses involving low failure probability levels. In this paper, the application of CNNs and the adaptive training strategy to slope reliability analysis in spatially variable soils is illustrated using a multi-layered ***c*-ϕ** slope case study, namely the Congress Street cut in Chicago (Oka & Wu 1990). The results are also compared against those published in several studies using the PCE-based methods to demonstrate the efficacy of the combination of CNNs and the adaptive training strategy.

2 METHODOLOGY AND IMPLEMENTATION

2.1 *Convolutional Neural Networks (CNNs)*

In the field of computer science, Convolutional Neural Networks (CNNs) specialize in processing data that has a grid-like topology, such as an image. This technique has been commonly used to solve many engineering problems (Xiao et al. 2018; Zhang et al. 2021b). In general, the major components of a CNN include, but are not limited to, an input layer, a convolutional layer, a pooling layer, an activation layer and an output layer.

In contrast to traditional neural networks that may not be effective in capturing the spatial information contained in an image, CNNs, through the use of convolutional filers, can extract features of the image taking into account pixel positions and the influence of nearby pixels. As a result, the spatial information can be better captured using CNNs than traditional neural networks. This attribute is particularly important for its adaptation to random field finite element analysis because spatial variability/spatial information is a key aspect of the information represented by a random field.

After extracting features contained in input images and passing them through a pooling layer, an activation layer and any other layer deemed appropriated for the task, a fully connected layer then links the extracted features to an output layer. Depending on the type of problem under study, a softmax layer or a regression layer can be specified for a classification task and a regression task, respectively. Therefore, the completion of an analysis using a CNN involves the combination of several types of layers. The configuration of these layers is often referred to as the architecture of a CNN. Details pertaining to these layers can be found in Cha et al. (2017).

2.2 *Interpreting random fields using CNNs*

In the context of random field finite element analysis (RF-FEM), the domain of a geosystem is first discretized into a system of grid points. Random fields that represent the spatial variation of a soil property are then generated in the soil domain of interest, following which each grid point in the soil domain is assigned a variable value of this soil property in accordance with the random field generated. Figure 1 shows samples of digitally generated random fields. The grid-like topology can be clearly seen in this figure.

A finite element model, on the other hand, needs also to be discretized into the finite element mesh, consisting of nodes and stress points. The integration of the two components then involves the mapping of the given realization of the random field to the stress points in the finite element mesh. Therefore, there is a one-to-one match between the grids of a random field and the mesh in a finite element model. In this manner, the finite element calculation can take into account the spatial variability contained in the generated random fields.

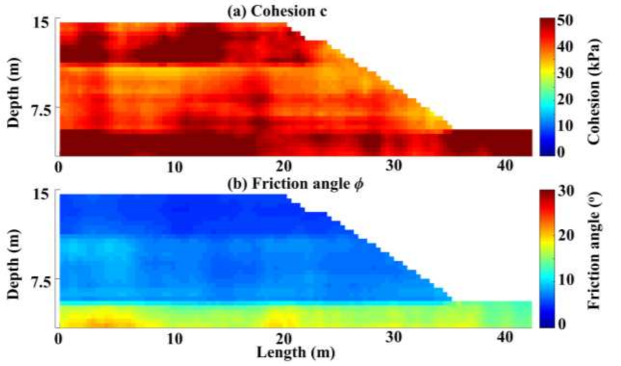

Figure 1. Samples of digitally generated random fields for (a) cohesion c and (b) friction angle ϕ.

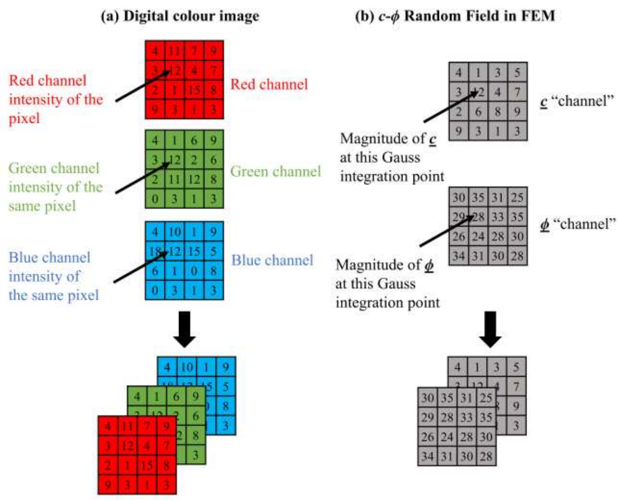

Figure 2. Comparison the representation of channel information in CNNs for (a) digital color images and (b) random fields in finite element models.

Given that random fields have a grid-like topology, it is hypothesized that CNNs are able to interpret and process random fields in a similar manner as they are used for images. The illustrative comparison of digital color images and random fields in Figure 2 verifies the hypothesis.

With reference to Figure 2(a), a color digital image is made of pixels. Each pixel has three channels that correspond to the three primary colors, i.e. red, green and blue. As a result, each pixel of the image is described by three values that correspond to the intensities of the three channels. In Figure 2(a), the input layer of CNNs contains an image of size 4 x 4 x 3. The two "4"s refer to the number of pixels that form the width and height of the image, while the "3" signifies that the image is made up of three channels, i.e. red, green and blue.

In random field finite element analysis, the stress integration points in the discretized finite element mesh or the grids in the discretized random field are the counterparts of the pixels in conventional image processing. The random field of a variable property then corresponds to the "channel". The magnitude of this variable property at a stress integration point/random field grid is analogous to the "channel intensity" of a "pixel".

With reference to the example shown in Figure 2(b), a soil layer characterized by a spatially varying cohesion $\boldsymbol{c}$ and friction angle $\boldsymbol{\phi}$ involves two random fields for each finite element realization. Accordingly, a two-channel "image", containing the "$\boldsymbol{c}$" channel and the "$\boldsymbol{\phi}$" channel, can be created to represent the spatial variability in the system. The values within each channel then correspond to the values of the spatially varying cohesion $\boldsymbol{c}$ and friction angle $\boldsymbol{\phi}$ respectively.

In this way, the two random fields are configured in a two-channel image-like manner for subsequent processing using CNNs. This is illustrated in Figure 2(b), in which the input random field image has a size of 4 x 4 x 2. The two "4"s refer to the number of stress integration points/random field grids along the width and height of the 2-D model domain, while the "2" indicates that there are two random fields involved, i.e. $\boldsymbol{c}$ and $\boldsymbol{\phi}$.

In summary, the comparison presented in Figure 2 highlights that there is a one-to-one match between the attributes of digital color images and random fields; therefore, CNNs should be able to interpret random fields in a similar manner as they are used for digital color images.

2.3 *Implementation procedure*

2.3.1 *Initialization*

With reference to Figure 3, after constructing the finite element model of the geotechnical system, X samples of random fields are generated following the statistics of the variable properties of interest. As will be explained in the later parts of this paper, the quantity of interest in the present study is the factor of safety (FoS) against the failure of the geotechnical system. In this regard, random field finite element analyses of the X samples are then carried out to obtain the associated factors of safety (FoS) of the geotechnical system under study.

The X pairs of random fields and their corresponding FoS values are then divided into a training set and a validation set. In the present study, 75% of the X dataset is used as the training set while the remaining 25% is used as the validation set.

After a CNN with an appropriate architecture is set up, the training and validation datasets are fed into the program for training. With sufficient training and proper validation, a CNN that gains the capability to predict the mapping function between input random fields and output FoS values can be obtained. This then concludes the initialization part of the implementation.

2.3.2 *Adaptive training*

As mentioned in the preceding section, the FoS against the failure of the geotechnical system is the key quantity of interest. A threshold FoS value of 1.0 is then used in the present study. Finite element simulations that yield a FoS value lower than this value are considered as failure cases. Therefore, the performance of the trained CNN for this region is of primary concern, and the adaptive training scheme is proposed to improve the accuracy of the trained CNNs over this region.

With refence to Figure 3, after a trained CNN is obtained at the end of the initialization phase, a set of random fields is generated. This is labelled as "Test set B" in Figure 3. The trained CNN is then used to predict the FoS values associated with this set of random fields, and a set of CNN-predicted FoS values can be obtained.

In the adaptive training strategy, a higher threshold FoS of 1.2 is used at this stage of the analysis to identify failure samples. Therefore, samples with CNN-predicted FoS values lower than 1.2 are considered as failure samples. Although a true failure sample is of a FoS value lower than 1.0, at this stage of the analysis, a relaxation in the threshold FoS value is adopted. The main reason behind this relaxation was as follows:

The "Test set B" is predicted using the trained CNN. Since the trained CNN is an approximate model of the finite element model, there is a risk that a sample with a CNN-predicted FoS value exceeding 1.0 is actually a failure case, meaning that the true FoS of this sample is less than 1.0. If the threshold were set to 1.0 as this stage, this failure sample would then be excluded from the adaptive sample set. In this regard, the relaxation of the threshold to 1.2 is introduced to reduce the risk that "true" failure samples would be falsely excluded.

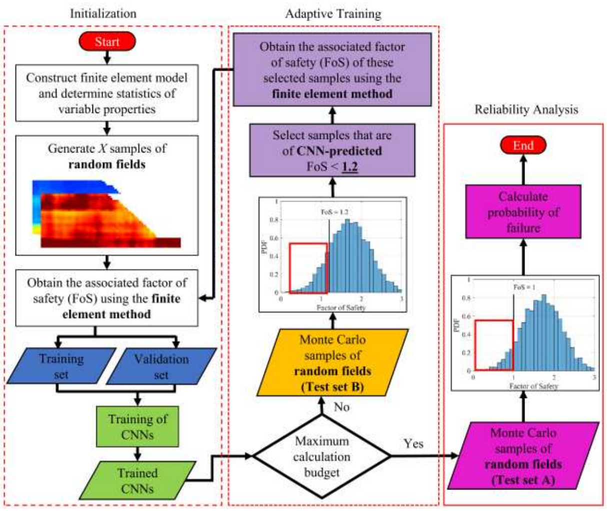

Figure 3. Implementation flow-chart for reliability analysis using CNNs.

The identified failure samples from "Test set B" are then fed back to the finite element model to obtain the true FoS values. In the next step, these samples and their associated true FoS values are combined with initial X samples to form a new dataset to re-train the CNN following the steps in the *Initialization* phase in Figure 3. In this way, the new training dataset contains more failure samples, which is expected to result in a trained CNN that can better predict the failure region.

The adaptive training phase in Figure 3 can be repeated so that additional failure samples can be identified for improved training. It is worth highlighting that predictions associated with "Test set B" are obtained using the trained CNN, and only those samples that are likely to be failure samples (determined using the trained CNN) are calculated using the finite element analysis. In this regard, the trained CNN acts as a filter that makes sure that the expensive finite element analysis is carried out only on samples that are deemed necessary and appropriate. Furthermore, the relaxed threshold FoS value of 1.2 is only used in the adaptive training phase. A threshold FoS value of 1.0 will be used for the subsequent reliability analysis.

2.3.3 *Reliability analysis*

After the CNN is sufficiently trained over the region of interest (FoS < 1.0) or the maximum computational budget is consumed, the trained CNN can be used to carry out reliability analysis following the Monte-Carlo simulation.

With reference to Figure 3, to perform the reliability analysis, a set of new random fields, labelled as "Test set A", is generated and predicted using the trained CNN. Samples with CNN-predicted FoS values lower than 1.0 are classified as failure samples. The probability of failure (P_{failure}) can then be calculated as $P_{failure} = size\ of\ failure\ samples\ /\ size\ of\ Test\ set\ A$.

3 CONGRESS STREET CUT, CHICAGO

3.1 *Case study description*

The Congress Street Cut is adopted as a case study to illustrate the capabilities of the proposed technique. Figure 4 shows that this slope consists of four layers. The spatial variabilities of the cohesion $\boldsymbol{c}$ and the friction angle $\boldsymbol{\phi}$ are considered in the three clay layers. The properties of the top sand layer are assumed to be deterministic. Table 1 summarizes the statistics of the random properties. In the present study, the cross-correlation between $\boldsymbol{c}$ and $\boldsymbol{\phi}$ is assumed to be zero for all three clay layers.

The finite element software Optum G2 (Krabbenhøft & Lyamin 2014) is used to perform the strength reduction calculations for evaluating the FoS values against slope failure. A built-in Karhunen-Loève (K-L) expansion with a single exponential autocorrelation function is used to generate random fields. In the present study, 1000 K-L expansion terms are used to ensure that the statistics are accurately represented by the generated random fields.

Figure 5 shows the architecture of the CNN used in the present study. The first layer contains input images of size 30 x 100 x 2. Then, a convolutional layer, consisting of 20 filters of size 25 x 90 with a padding of 0 and a stride of 1, is coupled with a max-pooling layer with a size of 2. Finally, a regression layer and a dropout layer are used to construct the regression between the FoS values and the extracted features. In the present study, a dropout rate of 50% is used to reduce overfitting.

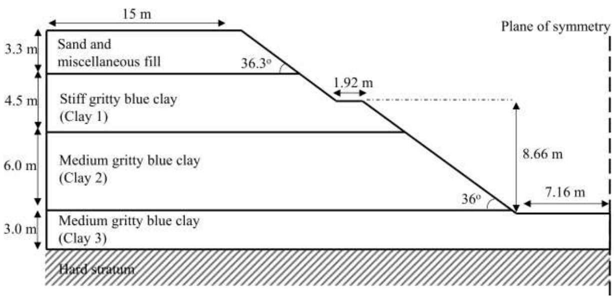

Figure 4. Geometry and geological profile of Congress Street cut.

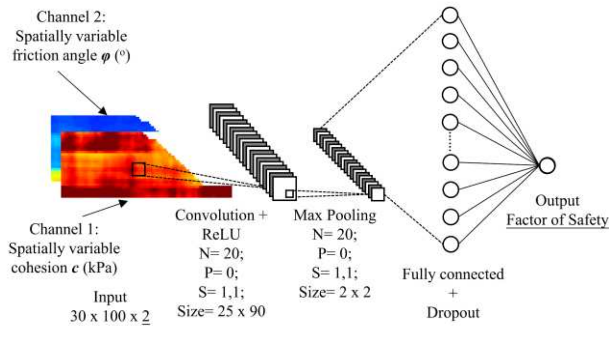

Figure 5. Architecture of CNNs. N = number of filters, P = padding, and S = stride.

3.2 *Results of reliability analysis*

Random field finite element analyses are carried out with 5000 Monte-Carlo simulations to provide the benchmark results. These 5000 samples correspond to the "Test set A" indicated in Figure 3. Based on the results of the 5000 finite element simulations, the true probability of failure (P_{failure}) is close to 0.0302. These values will be used to evaluate the accuracy of the CNN-predicted results in the subsequent parts of this paper.

With reference to Figure 3, the initialization phase requires X samples of random fields for training. An initial 100 random field samples are first generated for training the CNN. During the adaptive training phase, a "Test set B" containing 1000 samples is used to generate the CNN-predicted FoS values. For the investigation of the influence of training sample size, the number of training samples experimented in the present studies are 100, 160, 220, 280, 340, 400. For example, the case trained using 400 samples then contain 100 initial samples and 300 "failure" samples that are obtained using the adaptive training strategy.

Figure 6 plots the CNN-predicted FoS values against the FEM-predicted FoS values for the case trained using the initial 100 samples. There are 5000 scatters that correspond to the 5000 Monte-Carlo samples. In general, the CNN-predicted FoS values agree reasonably well with the FEM-predicted FoS values, which implies that CNN is a reasonable metamodel to map the function between input random fields and FoS values.

Figure 7 shows the FEM versus CNN cumulative probability distribution of FoS values for the set of data shown in Figure 6, using only the data in the failure region, where FoS <1.0. In contrast to Figure 6, Figure 7 implies that there are some major discrepancies between the FEM and CNN predicted results. As a result, the proposed adaptive training strategy is used to improve the performance of the CNN over the failure region.

Figure 8 shows, in a similar manner as in Figure 6, the results of the case with 400 training samples, which contain 100 initial samples and 300 "failure" samples obtained using the adaptive training strategy. The larger value of r^2 implies that the overall accuracy of the trained CNN has improved. In addition, improvements over the failure region are also observed, as can be seen from the reduced scatter in the zone where FoS < 1.0.

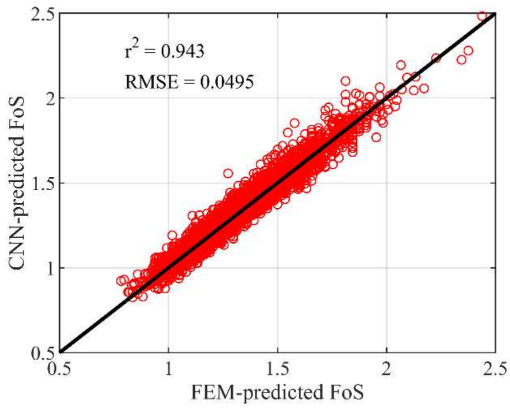

Figure 6. FEM versus CNN predictions of FoS for the case trained using 100 initial samples.

Table 1 Statistical parameters of the random properties (Jiang & Huang 2016)

Layers	Unit Weight (kN/m^3)	Cohesion c (kPa)			Friction angle ϕ (°)			Scale of fluctuation (SOF) (for both c and ϕ)	
		Mean	COV	Distribution	Mean	COV	Distribution	SOF_h (m)	SOF_v (m)
Sand	21	0	-	-	30	-	-	-	-
Clay 1	19.5	55	0.37	Lognormal	5	0.2	Lognormal	35.45	10.63
Clay 2	19.5	43	0.19	Lognormal	7	0.21	Lognormal	35.45	10.63
Clay 3	20	56	0.2	Lognormal	15	0.24	Lognormal	35.45	10.63

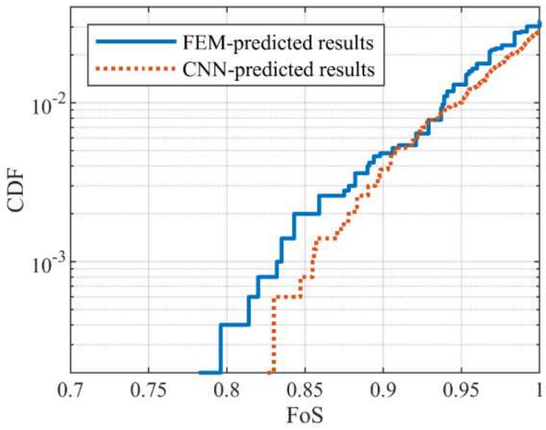

Figure 7. FEM versus CNN cumulative probability distribution of FoS for the case trained using 100 initial samples.

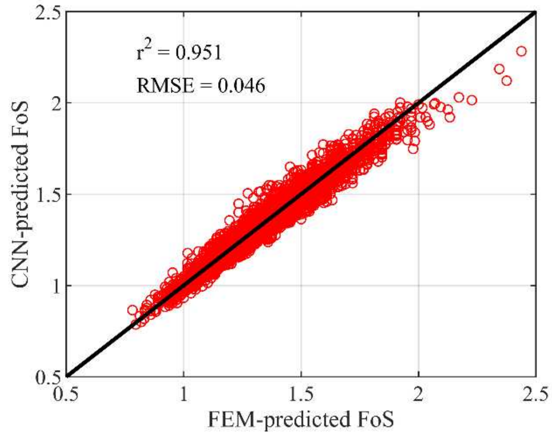

Figure 8. FEM versus CNN predictions of FoS for the case trained using 400 samples (100 initial samples and 300 "failure" samples)

Figure 9 shows, in a similar manner as in Figure 7, the FEM versus CNN cumulative probability distribution of FoS. In contrast to Figure 7, significant improvements in the accuracy of the trained CNN are observed. The discrepancies shown in Figure 7 are greatly reduced.

Figure 10 compares the CNN-calculated probability of failure ($P_{failure}$) for cases involving different training sample sizes. In addition, the results obtained using the adaptive training strategy are compared with those obtained without using the adaptive training strategy. The non-adaptive training strategy specifies that training samples are generated in a random manner. Furthermore, due to the stochastic nature of the analysis, the training of the CNNs are repeated 20 times following the recommendations in Wang & Goh (2021). The bars in Figure 10 represents the averaged predictions across the 20 sets of analyses while the error bars correspond to the maximum and minimum predicted values.

In general, both schemes perform well for predicting the failure probability. The percentage errors in the predicted $P_{failure}$ are all smaller than 20%. However, with the same number of training samples, the adaptive training strategy provides better results. For example, when 220 samples are used for training, the adaptive training strategy produces an error of about 2%, which contrasts to the 10% obtained in the non-adaptive training scheme. Furthermore, the results obtained using the adaptive training strategy also exhibit less variations. For example, with 220 training samples, the non-adaptive training scheme predicts errors ranged from 50% to -20% while the adaptive training strategy yields errors ranged from 20% to -15% only.

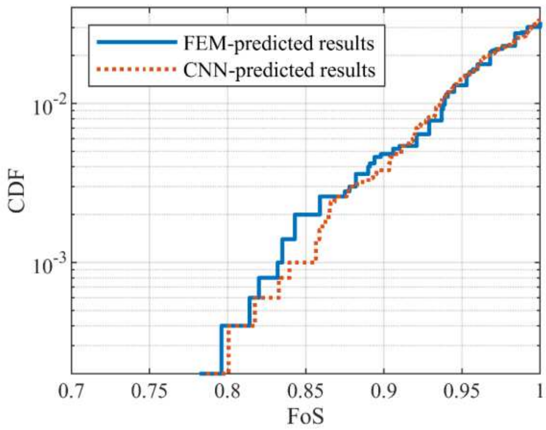

Figure 9. FEM versus CNN cumulative probability distribution of FoS for the case trained using 400 samples (100 initial samples and 300 "failure" samples).

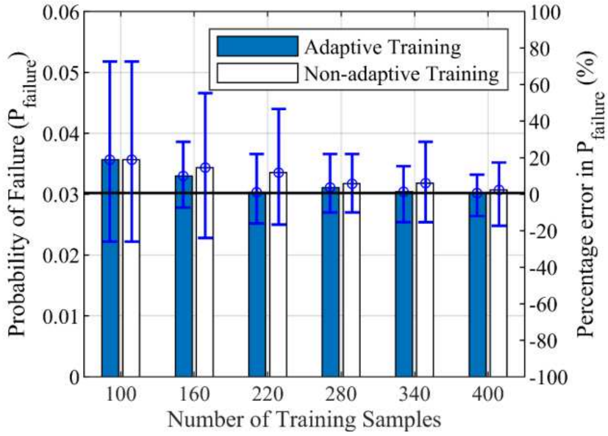

Figure 10. Comparison of CNN-calculated probability of failure ($P_{failure}$) between adaptive training and non-adaptive training.

Such observations are not unreasonable. The adaptive training strategy is designed to search for failure samples while the non-adaptive training scheme does not differentiate failure samples and non-failure samples. As a result, for a given size of training samples, the training sample set used for the adaptive training scheme contains more failure samples than the set used for the non-adaptive training scheme. It is then reasonable to have the adaptive training scheme better predicts over the failure region.

3.3 *Comparison with other techniques*

The probabilistic treatment of the Congress Street cut has been performed and documented in several publications. Therefore, in this section, the results of reliability analyses obtained using the trained CNN are compared against studies carried out by Jiang & Huang (2016); Li et al. (2016) and Li et al. (2019). In these studies, the Polynomial Chaos Expansion (PCE) method was used. Five variants of the PCE-based methods, labelled as Methods A to E in Figure 11, are compared with the CNN-based methods. The comparison is based on the percentage error in the metamodel-based value of $P_{failure}$. The error is calculated with respect to the benchmark results obtained using the direct Monte-Carlo simulation. The number of training samples used by the individual method is also indicated in the figure.

In general, the CNN-based methods are of better accuracy. The five variants of PCE-based methods are of error percentages ranged from -5% to -40% while the percentage errors of two variants of CNN-based method are both lower than 5%. In addition, Methods B to E require more than 1000 training samples, which can potentially hinder their application in real world situations. In contrast, the two variants of CNN-based methods yield improved accuracy with a significant reduction in the size of training samples, and CNN-based methods also outperform Method A, the most recent variant of the PCE-based method. In this regard, the proposed CNN-based metamodeling technique is a more accurate and efficient technique than the PCE-based methods.

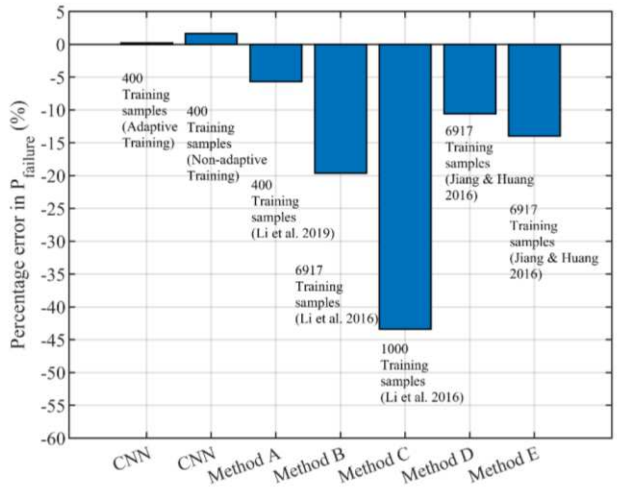

Figure 11. Comparison of the accuracy of CNN against different probabilistic metamodel-based techniques (Li et al. 2016; Jiang & Huang 2016; Li et al 2019; Wang & Goh 2021).

4 CONCLUSIONS

This paper presents the use of Convolutional Neural Networks as a metamodel for efficient slope reliability analysis in spatially variable soils. CNNs are capable of recognizing high-level features that contain information pertaining to the spatial variability of soil properties. With an appropriate training and validation, CNNs can successfully map the relationship between input random fields and the output factor of safety predictions. Therefore, CNNs can replace the computationally demanding random field finite element model for reliability analysis.

Furthermore, an adaptive training scheme is proposed to improve the performance of the CNN over the failure region. A comparison between the results obtained with and without the adaptive training strategy highlights that the proposed adaptive training strategy is effective in improving the accuracy of the CNN over the failure region.

Finally, a comparative study involving CNN-based methods and PCE-based methods, with the latter generally regarded as being the current state-of-the-art, shows that CNN-based methods are more accurate and efficient. In this regard, CNNs can offer a highly promising metamodel-based method for efficient reliability analysis in spatially variable soils.

5 ACKNOWLEDGEMENTS

This research was supported and conducted at the Centre for Protective Technology at the National University of Singapore. Special thanks to Dr. Numa Joy Bertola from École polytechnique fédérale de Lausanne for drafting the abstract in French.

6 REFERENCES

Abdel-Hamid, O., Mohamed, A. R., Jiang, H., Deng, L., Penn, G., & Yu, D. (2014). Convolutional neural networks for speech recognition. *IEEE Transactions on Audio, Speech and Language Processing*. https://doi.org/10.1109/TASLP.2014.2339736

Ang, A. H. S., & Tang, W. H. (2007). *Probability concepts in engineering planning and design: Emphasis on application to civil and environmental engineering*. Wiley.

Au, S. K., & Beck, J. L. (2001). Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics*. https://doi.org/10.1016/S0266-8920(01)00019-4

Cha, Y. J., Choi, W., & Büyüköztürk, O. (2017). Deep Learning-Based Crack Damage Detection Using Convolutional Neural Networks. *Computer-Aided Civil and Infrastructure Engineering*. https://doi.org/10.1111/mice.12263

Gao, G. H., Li, D. Q., Cao, Z. J., Wang, Y., & Zhang, L. (2019). Full probabilistic design of earth retaining structures using generalized subset simulation. *Computers and Geotechnics*. https://doi.org/10.1016/j.compgeo.2019.04.020

Goh, A. T. C., Zhang, W. G., & Wong, K. S. (2019). Deterministic and reliability analysis of basal heave stability for excavation in spatial variable soils. *Computers and Geotechnics*. https://doi.org/10.1016/j.compgeo.2018.12.015

Griffiths, D. V., Huang, J., & Fenton, G. A. (2009). Influence of spatial variability on slope reliability using 2-D random fields. *Journal of Geotechnical and Geoenvironmental Engineering*. https://doi.org/10.1061/(ASCE)GT.1943-5606.0000099

Jiang, S. H., & Huang, J. S. (2016). Efficient slope reliability analysis at low-probability levels in spatially variable soils. *Computers and Geotechnics*. https://doi.org/10.1016/j.compgeo.2016.01.016

Krabbenhøft K, Lyamin AV. Optum G2. Optum Computational Engineering. 2014.

Li, D. Q., Zheng, D., Cao, Z. J., Tang, X. S., & Phoon, K. K. (2016). Response surface methods for slope reliability analysis: Review and comparison. *Engineering Geology*. https://doi.org/10.1016/j.enggeo.2015.09.003

Li, D. Q., Zheng, D., Cao, Z. J., Tang, X. S., & Qi, X. H. (2019). Two-stage dimension reduction method for meta-model based slope reliability analysis in spatially variable soils. *Structural Safety*. https://doi.org/10.1016/j.strusafe.2019.101872

Liu, W. S., Cheung, S. H., & Cao, W. J. (2019). An efficient surrogate-aided importance sampling framework for reliability analysis. *Advances in Engineering Software*. https://doi.org/10.1016/j.advengsoft.2019.102687

Oka, Y., & Wu, T. H. (1990). System reliability of slope stability. *Journal of Geotechnical Engineering*. https://doi.org/10.1061/(ASCE)0733-9410(1990)116:8(1185)

Vanmarcke, E. (2010). *Random fields: analysis and synthesis*. World scientific.

Wang, L., Wu, C., Tang, L., Zhang, W., Lacasse, S., Liu, H., & Gao, L. (2020). Efficient reliability analysis of earth dam slope stability using extreme gradient boosting method. *Acta Geotechnica*. https://doi.org/10.1007/s11440-020-00962-4

Wang, Z.Z., & Goh, S. H. (2021). Novel approach to efficient slope reliability analysis in spatially variable soils. *Engineering Geology*, *281*, 105989. https://doi.org/10.1016/j.enggeo.2020.105989

Wang, Z. Z., Xiao, C., Goh, S. H., & Deng, M.-X. (2021). Metamodel-Based Reliability Analysis in Spatially Variable Soils Using Convolutional Neural Networks. *Journal of Geotechnical and Geoenvironmental Engineering*. https://doi.org/10.1061/(asce)gt.1943-5606.0002486

Xiao, C., Qin, R., Huang, X., & Li, J. (2018). A STUDY of USING FULLY CONVOLUTIONAL NETWORK for TREETOP DETECTION on REMOTE SENSING DATA. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. https://doi.org/10.5194/isprs-annals-IV-1-163-2018

Zhang, J. Z., Huang, H. W., Zhang, D. M., Zhou, M. L., Tang, C., & Liu, D. J. (2021a). Effect of ground surface surcharge on deformational performance of tunnel in spatially variable soil. *Computers and Geotechnics*. https://doi.org/10.1016/j.compgeo.2021.104229

Zhang, J. Z., Phoon, K.K., Zhang, D. M., Huang, H.W., & Tang, C. (2021b). Novel approach to estimate vertical scale of fluctuation based on CPT data using convolutional neural networks. *Engineering Geology*. https://doi.org/10.1016/j.enggeo.2021.106342.