Journal of Building Engineering 124 (2026) 115936

Contents lists available at ScienceDirect

Journal of Building Engineering

journal homepage: www.elsevier.com/locate/jobe

Automated design of reinforced concrete dapped-end connections using hybrid deep learning and generative AI augmentation

Shashika Dharmawansha [a,b], Sumudu Herath [b], P.L.N. Fernando [b], D.P.P. Meddage [c,d,*], Chathura Rajapakse [b]

[a] Department of Engineering, University of Cambridge, Cambridge, CB2 1PZ, United Kingdom
[b] Department of Civil Engineering, University of Moratuwa, Sri Lanka
[c] School of Engineering and Information Technology, The University of New South Wales, Canberra, ACT, 2600, Australia
[d] Department of Civil Engineering, Faculty of Engineering, Sri Lanka Institute of Information Technology, Sri Lanka

ARTICLE INFO

Keywords:
Dapped-end connections
Predictive modelling
Orthogonal reinforcement layout
Deep learning
Generative adversarial networks

ABSTRACT

Dapped-end connections, also known as half-joints or Gerber beams, are widely used yet structurally vulnerable elements in precast concrete structures due to high stress concentrations near the re-entrant corner. Therefore, a comprehensive assessment of the load-bearing capacity of dapped-end connections is important to ensure structural integrity and mitigate the risk of failure. Although prior studies have explored their behaviour through analytical and experimental methods, the application of data-driven approaches remains limited due to the availability of limited experimental data, which constrains the predictive accuracy and generalisation of Machine Learning (ML) models. This study presents a novel approach that integrates numerical simulation with Conditional Tabular Generative Adversarial Network (CTGAN)-based data augmentation to enhance prediction accuracy and model generalisation. A numerical database containing 720 results was developed, which was expanded with 680 augmented data using CTGAN. The combined dataset of 1400 instances was used to train Artificial Neural Network (ANN), Genetic Algorithm-ANN (GA-ANN), and Particle Swarm Optimisation-ANN (PSO-ANN) models. The hybrid models outperformed the standalone ANN, with GA-ANN achieving the highest accuracy (testing $R^2 = 0.961$). The trained models were separately validated using 64 unseen experimental datasets, which shows the improved generalisation of the models through augmentation. Shapley Additive Explanations analysis reveals that the GA-ANN model predictions aligned with the principles underlying the compatibility of deformations of dapped-ends. Further, a novel ML-assisted design model was developed, which predicts multiple solutions for a given design problem, assisting in the optimisation of connection design.

1. Introduction

Reinforced concrete dapped-end connections, also termed half-joints or Gerber beams, are widely implemented in precast bridges, residential and office complexes, industrial buildings, and parking structures. Their characteristic shape, with a sudden reduction in depth at a sharp re-entrant corner, makes them useful for connecting precast concrete structural elements while preserving a reduced

* Corresponding author. School of Engineering and Information Technology, The University of New South Wales, Canberra, ACT, 2600, Australia.
E-mail addresses: p.meddage@unsw.edu.au, pasindu95dm@gmail.com (D.P.P. Meddage).

https://doi.org/10.1016/j.jobe.2026.115936
Received 31 October 2025; Received in revised form 12 March 2026; Accepted 21 March 2026
Available online 3 April 2026

structural depth. However, this shape leads to concentration of stresses at the re-entrant corner, which can result in cracking under service loads [1,2]. These connections transmit high shear forces through considerably reduced sections, increasing their susceptibility to brittle shear failures. Furthermore, durability issues such as corrosion caused by the ingress of water and salts into the cracks make dapped-end connections particularly challenging to design, maintain and assess [1].

Over the past decades, experimental and analytical studies have been carried out to investigate the behaviour of dapped-end connections. The analytical studies include Strut-and-Tie Models (STM) [3–5], stress-field analysis [6,7], empirical and semi-empirical methods [8], numerical analysis using nonlinear finite element methods [9–14] and kinematics-based methods [15,16]. Dapped-end specimens have been experimentally investigated by varying geometric and material properties [1,6,10,17–35]. These studies have built a strong foundation for understanding load transfer mechanisms, identifying critical detailing requirements and supporting safe design.

With recent developments in Machine Learning (ML), the research community has increasingly adopted data-driven frameworks capable of estimating structural capacity directly from geometric and material parameters. Even though these approaches are widely used in the concrete domain [36–40], their application to reinforced concrete dapped-end connections remains limited, with only a few studies in the literature addressing the prediction of dapped-end connection capacity using machine learning approaches [41]. While accurate prediction of the capacity of dapped-end connections for given design variables, including reinforcement layout, material properties and geometry, remains essential, additional benefits can be achieved through the development of an ML-assisted design framework capable of identifying suitable reinforcement configurations that satisfy a target performance requirement. Such an approach enables systematic exploration of the design space and provides a complementary tool for design-oriented decision-making when multiple interacting parameters govern the structural response.

Therefore, this study reports a novel ML-assisted design framework for reinforced concrete dapped-end connections by integrating data augmentation with deep neural networks. A numerical database is developed through a parametric study and subsequently used for data augmentation using a generative adversarial network to generate statistically consistent synthetic samples. The combined numerical and augmented dataset is then employed to train and validate machine learning models for predicting the ultimate capacity of dapped-end connections. Further validation is carried out by assessing the predictive capability of the developed models against unseen experimental data available in the literature. Moreover, the conventional forward prediction pipeline is extended by developing a reverse prediction framework to identify the input geometries, concrete compressive strength and reinforcement layouts required to achieve a specified target capacity, which is particularly useful for practical design applications.

1.1. Related work

This study primarily focuses on numerical modelling, data augmentation and the application of machine learning for developing an ML-assisted design framework for reinforced concrete dapped-end connections. Numerical analyses are widely used to investigate the behaviour of dapped-end connections [1,11–13], which provide a cost-effective and time-efficient alternative to experimental testing while offering greater flexibility in parametric studies. Following their experimental study, Rajapakse et al. [1] used validated numerical models to examine how different reinforcement arrangements influence internal force transfer and stress distribution in dapped-end connections. Aksoylu et al. [12] conducted both experimental and numerical investigations of shear strength at dapped-end beams having different shear spans, where a numerical parametric study was conducted to further examine the effects of the span-to-depth ratio and the ratio of the distance of the load on the beam from the support and the depth of the beam. Santarsiero and Picciano [13] conducted a numerical investigation on post-tension retrofitting of reinforced concrete dapped-end beams using validated numerical models against experimental test results of Rajapakse et al. [1].

Later, various research communities have used data-driven methods such as ML in the domain of reinforced concrete structures [36–39]. However, a very limited number of ML-related research studies are available on dapped-end connections. Picciano et al. [41] utilised a database of 210 experimental data instances for both orthogonal and diagonal reinforcement layouts from 22 published experimental studies to train regression ML models to predict the ultimate capacity of dapped-end connections. Among the trained models, the Gaussian Process Regressor exhibited the highest accuracy, achieving a training R^2 score of 0.958 and a testing R^2 score of 0.984. While accurate prediction of the capacity of dapped-end connections remains essential, further benefits can be gained from the development of an ML-assisted design model to predict the suitable reinforcement configurations to achieve a user-given dapped-end capacity. The reverse-engineered workflow of trained ML models enables this functionality and is particularly valuable in civil engineering applications. However, the primary limitation in developing more generalisable models lies in the limited availability of experimental data.

Recently, data augmentation techniques have been widely adopted to enhance the databases and improve the performance of ML predictions [42]. Among these techniques, more advanced approaches such as Generative Adversarial Networks (GANs) [43] have demonstrated strong potential for generating realistic synthetic data across a wide range of applications [39,42,44,45]. Previous studies have shown that augmenting datasets with GAN-generated synthetic data consistently improves both prediction accuracy [46–50] and generalisation capability of ML models when predicting unseen experimental data [46,47]. For instance, Chen et al. [45] applied GAN-generated data to model the compressive strength of concrete with industrial waste. The model based on the original (augmented) data achieved an R^2 of 0.89 (0.98). Similarly, Marani et al. [46] reported that the best-performing ML model achieved an R^2 score of 0.96 in predicting unseen experimental results, highlighting the effectiveness of synthetic data in enhancing model generalisation. Accordingly, GAN-based data augmentation enables the generation of realistic and diverse training samples, which improves both prediction accuracy and generalisation by providing a more representative training distribution. Conditional Tabular Generative Adversarial Network (CTGAN) [51] is an extension of GAN specifically developed for tabular data, which has the ability to

generate both continuous and discrete data using conditional sampling and mode-specific normalisation.

1.2. Methodology

This study utilised a novel approach that combines numerical simulation with CTGAN-based data augmentation to improve both predictive accuracy and model generalisation of hybrid neural networks in predicting the capacities of dapped-end connections with an orthogonal reinforcement layout. The research process involved several steps, as illustrated in Fig. 1. First, a total of 720 numerical simulations were performed to create a numerical database following a comprehensive model validation for 24 experimental tests. The database was expanded with 680 synthetic data samples generated using CTGAN. Additionally, a benchmark ANN was used to ensure the consistency of feature distributions. Before training the current ANN models, the dataset was rechecked for feature distribution (using mean and standard deviation) and remaining outliers to ensure data quality and consistency. To evaluate the effect of augmentation on generalisation, two separate training cases were considered. Model 1 was trained using the numerical database only (720 instances). Model 2 was trained using the combined database (1400 instances). For each training case, ANN, GA-ANN and PSO-ANN models were developed. The generalisation capability of the trained models was assessed using unseen experimental data (64 experimental results).

SHapley Additive exPlanations (SHAP) [52] was employed to explain how input parameters (such as material properties, geometric parameters, and reinforcement configuration) influence the capacity of dapped-end connections and to verify whether the ML model has captured the underlying engineering principles governing their behaviour. Significance of this research extends beyond accurately predicting the capacities of dapped end connections with an orthogonal reinforcement layout. It demonstrates the potential of data augmentation techniques, specifically the CTGAN, in enhancing the ML performance. Lastly, an ML-assisted predictive modelling tool was developed to predict the different combinations of design parameters of dapped-end connections required to achieve user-specified capacity within a given error margin. This tool is particularly valuable for assessing and optimising dapped-end connections in civil engineering applications.

2. Numerical modelling and generative modelling

This section presents the development of a comprehensive database consisting of 1400 data instances, generated through a combination of numerical simulations and data augmentation using CTGAN. The numerical database was developed following a thorough model validation against 24 experimental results to ensure its accuracy. CTGAN was then employed to generate realistic synthetic data, aiming to enhance the database and improve the performance of the ML models.

2.1. Input parameter selection

The main parameters contributing to the strength of dapped-end connections were selected as input parameters for ML model

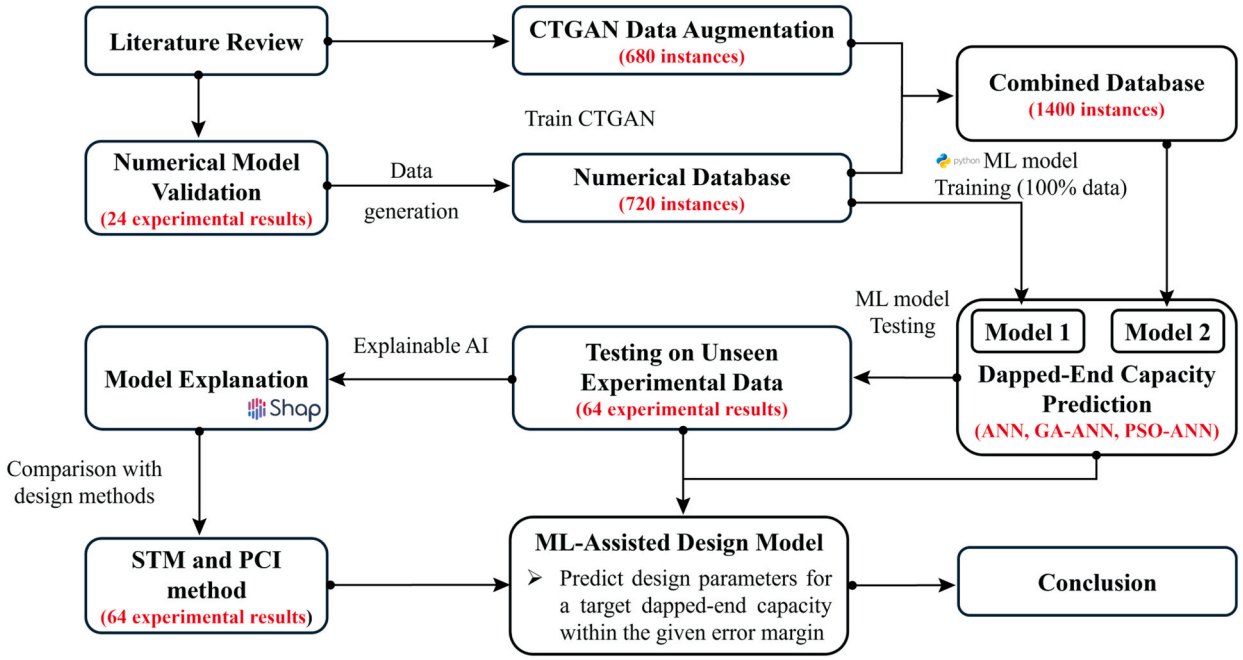


Fig. 1. Workflow of the study.

training to predict the ultimate capacity. The selection criteria for the input parameters were based on the behaviour and load-transfer mechanism of dapped-end connections, which can be visualised with the aid of STMs. A typical reinforcement arrangement of dapped-end connections with an orthogonal reinforcement layout, along with a STM [3] is illustrated in Fig. 2. STM features main horizontal and vertical ties at the re-entrant corner (reinforcement in tension), as well as an inclined strut in the dapped end (concrete in compression). The nib transfers vertical forces through an inclined compressive concrete strut, as shown in Fig. 2. To maintain equilibrium, the presence of vertical ties, provided by vertical stirrups (A_{sv}) and horizontal ties, provided by horizontal reinforcement (A_{sh}) are required.

The selected main parameters influencing this mechanism include the dapped-end depth (h), which governs the size and inclination of the strut and the shear-span-to-effective-depth ratio (a/d), which affects the distribution of force. The compressive strength of concrete (f_c) influences the load-bearing capacity of the strut, while the strength of flexural ($A_{sh}\ f_{y,sh}$) and hanger ($A_{sv}\ f_{y,sv}$) reinforcements represent the capacity of the horizontal and vertical ties, respectively, where $f_{y,sh}$ and $f_{y,sv}$ denote the yield strength of flexural and hander rebars, respectively. The number of flexural (N_h) and hanger (N_v) reinforcement layers, along with the spacing between flexural (S_h) and hanger (S_v) reinforcement layers, were selected to account for the effect of reinforcement distribution and to allow the ML-assisted design model to generate multiple reinforcement configurations to achieve the target dapped-end capacity.

2.2. Numerical modelling of dapped-end connections

This section describes the development of a numerical database through nonlinear Finite Element (FE) analyses using VecTor2 software [53] based on experimental investigations conducted by Mata Falcon et al. [6] and Rajapakse et al. [1]. The applied boundary conditions along with the material models used are illustrated in Fig. 3 where only half of the beam is modelled, considering the symmetric nature of the simulated event and the computational efficiency [1,54]. VecTor2 operates based on the Disturbed Stress Field Model (DSFM) [55], which is a smeared rotating crack approach derived from the Modified Compression Field Theory (MCFT) for reinforced concrete elements subjected to shear.

To ensure that the FE analyses can be reproduced conveniently, effects such as tension stiffening, aggregate interlock, reinforcement yielding and concrete softening were incorporated using the default advanced relationships implemented in VecTor2 with a few exceptions related to the modelling of concrete. The pre-peak constitutive behaviour of concrete in compression was modelled using the Popovics Normal Strength Concrete model, while compression softening was represented by the Vecchio 1992-A (e1/e2-Form) formulation. The steel reinforcement was modelled with a bilinear stress–strain relationship that includes strain hardening. Quadrilateral plane stress elements were employed to model concrete and discrete truss elements were used to represent reinforcement under the assumption of a perfect bond. The analyses were performed under displacement-controlled loading, where 15 mm displacement was applied progressively in incremental steps of 0.01 mm until failure.

A total of 24 test specimens were considered for the validation of the numerical model, where 16 test specimens for a 300 mm dapped-end depth (DEB-1.1 – DEB-1.9 specimens) were taken from Mata Falcone et al. [6] test specimens (OL-1 – OL-8 specimens) for a 500 mm dapped-end depth were taken from the experimental study of Rajapakse et al. [1]. Test specimens from two different experimental studies [1,6] with two beam sizes representing both flexural and shear failures, were selected to enhance the robustness of numerical model validation. Table 1 presents the model validation for predicting the ultimate capacity of dapped-end connections using VecTor2.

Fig. 4 (a) presents a summary of the comparison between the ultimate capacities obtained from the experimental and FE analysis. For beams with a 300 mm dapped-end depth, the average ratio of experimental to predicted ultimate capacity ($V_{u,exp}/V_{u,FEA}$) is 0.97 and the Coefficient of Variation (COV) is 6.95%. Similarly, for beams with a 500 mm dapped-end depth, the corresponding values are $V_{u,exp}/V_{u,FEA}$ of 0.99 and COV of 4.27%. Fig. 4 (a) highlights three distinct crack patterns corresponding to lightly (OL-1), moderately

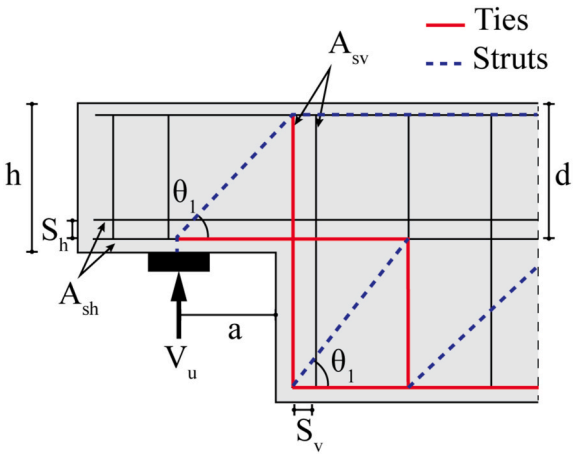
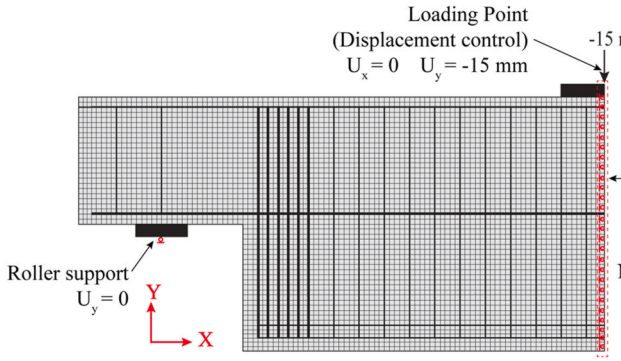


Fig. 2. Typical orthogonal reinforcement arrangement along with STM.



Concrete Models	
Compression pre-peak	Popovics (NSC)
Compression post-peak	Popovics/Mander
Compression softening	Vecchio 1992-A (e1/e2 –form)
Crack width check	Agg/2.5 Max Crack width
Mesh Details	
Element type (Concrete)	Four node rectangular element
Element type (Rebar)	Two-node truss bar element
Element size (Concrete)	20 mm x 20 mm
Maximum aspect ratio	1.5
Loading Details	
Type	Displacement control
Load increment factor	0.01
Number of load stages	101
Maximum number of iterations	100

Fig. 3. Summary of the finite element modelling approach.

Table 1
Comparison of experimental and FE-predicted ultimate capacities of dapped-end connections.

Literature	Dapped-end depth - *h* (mm)	Specimen name	$V_{u,Exp}$ (kN)	$V_{u,FEA}$ (kN)	$V_{u,Exp}/V_{u,FEA}$
Mata Falcon et al. [6]	300	DEB 1.1 (T1)	193.6	194.0	0.99
		DEB.1.2 (T1)	145.8	136.9	1.07
		DEB.1.2 (T2)	132.7	136.9	0.97
		DEB.1.3 (T1)	121.1	134.1	0.90
		DEB.1.3 (T2)	133.0	134.1	0.99
		DEB.1.4 (T1)	183.0	189.0	0.97
		DEB.1.4 (T2)	170.4	189.0	0.90
		DEB.1.5 (T1)	125.3	132.1	0.95
		DEB.1.6 (T1)	309.2	257.8	1.19
		DEB.1.6 (T2)	250.9	257.8	0.97
		DEB.1.7 (T1)	194.4	203.1	0.96
		DEB.1.7 (T2)	188.8	203.1	0.93
		DEB.1.8 (T1)	195.3	206.4	0.95
		DEB.1.8 (T2)	199.1	206.4	0.96
		DEB.1.9 (T1)	141.7	152.7	0.93
		DEB.1.9 (T2)	145.5	152.7	0.95
				Mean	0.97
				COV (%)	6.95
Rajapakse et al. [1]	500	OL-1	245	255.6	0.96
		OL-2	283	308.3	0.92
		OL-3	472	458.1	1.03
		OL-4	555	559.2	0.99
		OL-5	628	663.1	0.95
		OL-6	728	741.7	0.98
		OL-7	868	830.1	1.05
		OL-8	995	965.1	1.03
				Mean	0.99
				COV (%)	4.27

(OL-4), and densely (OL-8) reinforced dapped-end connections under different failure modes. The $V_{u,exp}/V_{u,FEA}$ ratios for these test results are 0.96, 0.99, and 1.03, respectively. This demonstrates consistent predictive accuracy of the numerical modelling approach and, confirms its unbiased performance in predicting ultimate capacities for dapped-end connections of varying sizes, reinforcement configurations, and failure mechanisms. The overall validation shows excellent agreement, with an average $V_{u,exp}/V_{u,FEA}$ of 0.98 and COV of 6.22% across all 24 test results.

Fig. 4 (b) presents a comparison of failure modes in lightly and densely reinforced dapped-end connections. The OL-1 specimen exhibits a flexural failure mode in lightly reinforced connections, characterised by a propagating diagonal corner crack. In contrast, the OL-8 specimen demonstrates a shear failure mode in densely reinforced connections, where a diagonal crack initiates from the support. This comparison confirms that the numerical model effectively captures both flexural and shear failure mechanisms along with crack patterns, further reinforcing the capability of VecTor2 FE models to simulate the behaviour of dapped-end connections across varying reinforcement configurations.

Fig. 5 compares the cover spalling behaviour observed in the experiment with the flow of compressive forces obtained from FEA for a dapped-end specimen with a 500 mm depth [1]. In the marked region, the flow of compressive forces is concentrated and deviated. This change in the direction of concentrated compression forces generates significant tensile stresses in the unreinforced concrete cover. If the induced stress exceeds the tensile strength of the concrete, cracks will propagate, ultimately causing the outer layer to

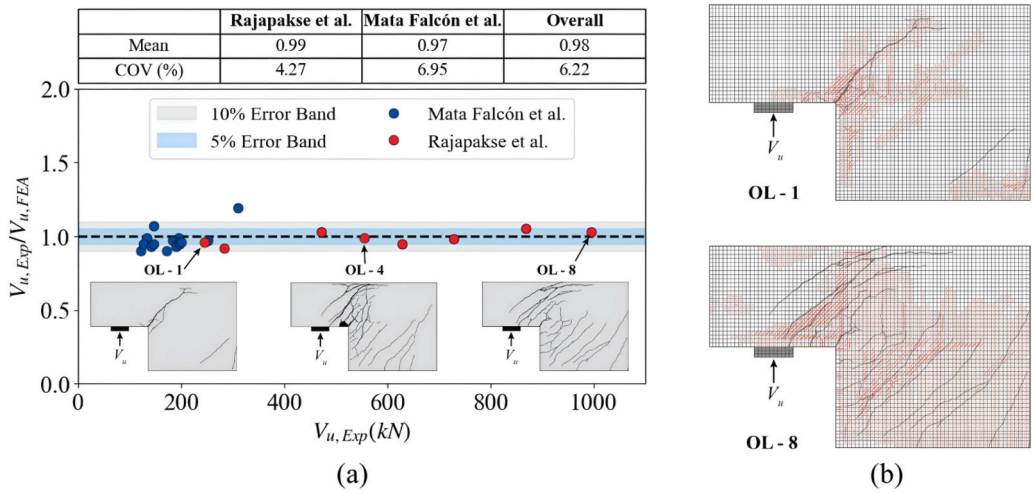


Fig. 4. Comparison of experimental and numerical predictions of dapped-end connections: (a) ultimate capacity and (b) comparison of experimental [1] and predicted failure modes.

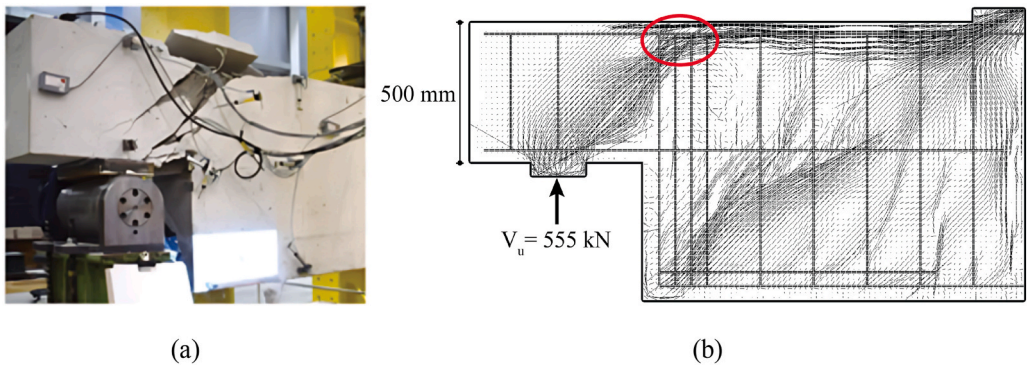


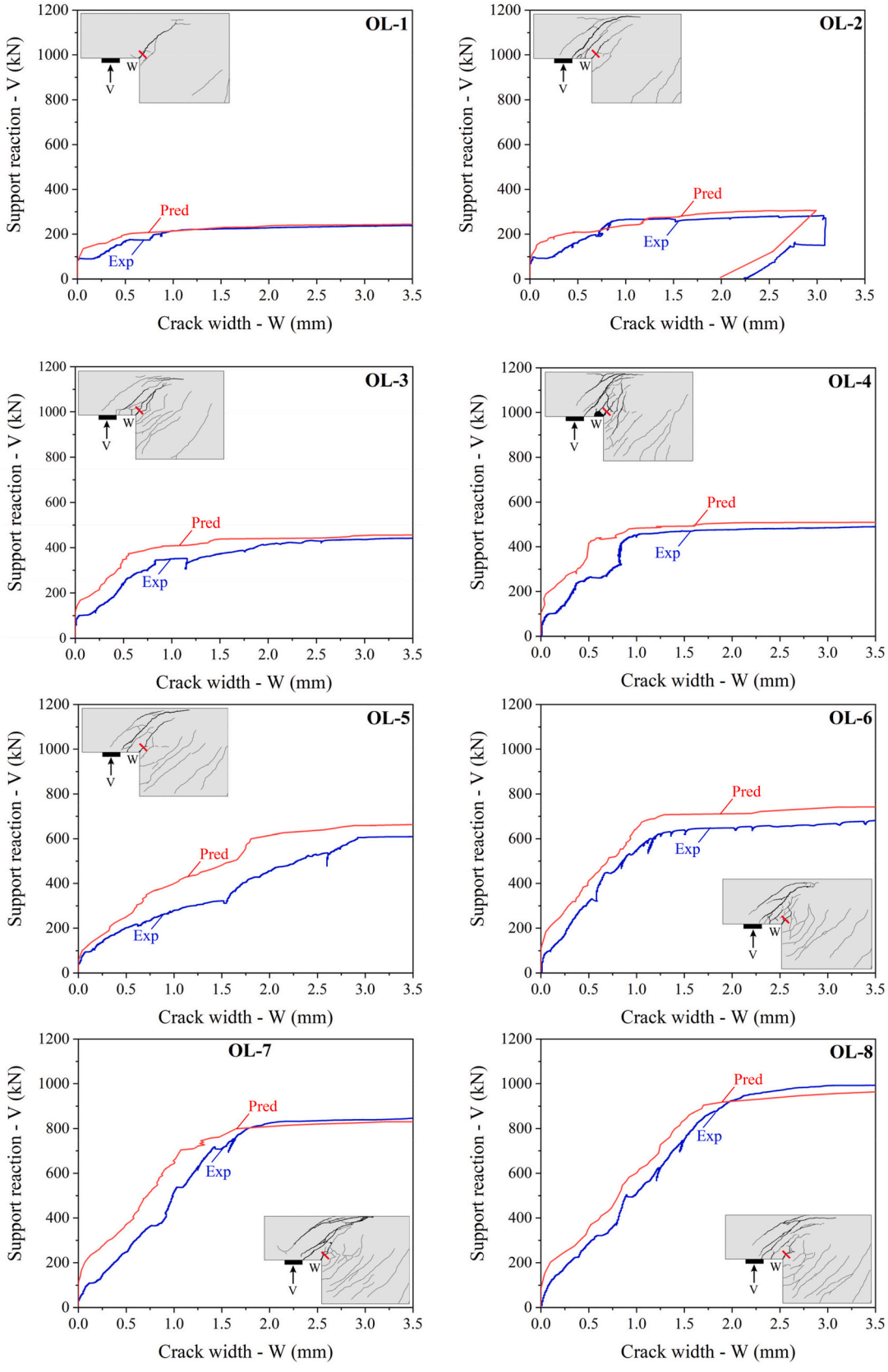
Fig. 5. Comparison of the cover spalling behaviour of dapped-end connections: (a) experimental response [1] and (b) predicted flow of compressive stresses (OL-4).

break away. This phenomenon is known as cover spalling. Therefore, this comparison shows that the FE model has identified such behaviours of dapped-end connections as well.

To further evaluate the reliability of the numerical modelling approach in capturing the experimental behaviour of dapped-end connections, the variation of support shear force (V) with corner crack width (W) was compared as shown in Fig. 6. The comparison indicates that the FE models have captured the variation between support shear force and corner crack width reasonably well. While the numerical models slightly underpredict the crack widths for a given support reaction, the overall trend is well captured across a spectrum of failure modes ranging from flexural failure in lightly reinforced connections to shear failure in densely reinforced specimens. The slight underprediction of crack width for a given support reaction value may be attributed to material model assumptions used in the numerical simulations, especially the consideration of a perfect bond between concrete and steel reinforcement.

2.2.1. Development of the numerical database

A comprehensive parametric study was conducted for beams with dapped-end depths of 200, 300, 400, 500 and 600 mm to generate a numerical database. The varied parameters include dapped-end depth (h), span to effective depth ratio (a/d), compressive strength of concrete (f_c), number of horizontal (N_h) and vertical (N_v) reinforcement layers, spacing between horizontal (S_h) and vertical (S_v) reinforcement layers and strength of horizontal ($A_{sh}f_{y,sh}$) and vertical ($A_{sv}f_{y,sv}$) reinforcements. The numerical database captures a range of variables designed to capture both primary shear and flexural failure modes in dapped-end connections. A total of 720 numerical results were generated which consists of 144 numerical results for each dapped-end depth. Table 2 summarises the combinations of input parameters considered in the numerical database. The selected parameter ranges were defined to cover a broad spectrum of practical variations, allowing the subsequent CTGAN-based data augmentation to effectively fill gaps in the continuous variables. Commercially available reinforcement diameters were selected to ensure practical relevance and consistency with common construction practice.



(caption on next page)

Fig. 6. Comparison of experimental (Rajapakse et al. [1]) and FE-predicted variation of support shear force with corner crack width.

Both single-layer and double-layer configurations were considered for the main horizontal (flexural-A_{sh}) reinforcement to capture the effect of reinforcement distribution and multiple layers were used for the main vertical (hanger-A_{vh}) reinforcement. This approach reflects real-world practice, where multiple hanger layers are typically provided to enhance load transfer. Furthermore, both four-legged and two-legged stirrups were employed for the hanger reinforcement to represent common detailing methods observed in practical construction scenarios.

As the a/d ratio is a critical parameter influencing the failure mechanism of dapped-end connections, lower a/d ratios are typically associated with shear-dominant behaviour, while higher ratios tend to show flexure-shear or flexural failure [29]. Therefore, nine a/d ratios ranging from 0.5 to 1.25 were considered to capture their influence on the structural response of dapped-end connections. These variations in the parametric study enable the representation of realistic dapped-end connection design conditions, allowing the use of generalised ML models capable of accurately predicting the actual capacities of dapped-end connections (see Section 4.3).

2.3. Conditional tabular generative adversarial networks (CTGAN)

Generative Adversarial Networks (GANs) which was introduced by Goodfellow et al. [43] belong to deep learning models that uses zero-sum game strategy and enable the learning of data distributions through an adversarial learning process. GAN consists of two interconnected neural networks, which are the Generator (G) and the Discriminator (D). The primary function of the generator is to learn the underlying data distribution and produce synthetic samples that closely imitate the real data, aiming to "fool" the

Table 2
Summary of the numerical database.

h (mm)	a/d	f_c (MPa)	N_h	S_h (mm)	φ_h (mm)	N_v	S_v (mm)	φ_v (mm)	No. of FE models
200	0.5	25, 60	1	0	10, 12	1	0	8, 10	8
			1	0	10, 12	2	40, 80	8, 10	16
			2	40, 80	10, 12	1	0	8, 10	16
			2	40, 80	10, 12	2	40, 80	8, 10	32
	0.9	30, 45	1	0	16, 20	3	30, 60	10, 12	16
			2	60, 100	16, 20	3	30, 60	10, 12	32
			1	0	16, 20	4	60	10, 12	8
			2	60, 100	16, 20	4	60	10, 12	16
300	0.8	25, 60	1	0	10, 16	1	0	8, 12	8
			1	0	10, 16	2	90, 180	8, 12	16
			2	40, 110	10, 16	1	0	8, 12	16
			2	40, 110	10, 16	2	90, 180	8, 12	32
	1	30, 45	1	0	20, 25	4	30, 50	8, 10	16
			2	60, 120	20, 25	4	30, 50	8, 10	32
			1	0	20, 25	5	50	8, 10	8
			2	60, 120	20, 25	5	50	8, 10	16
400	0.6	25, 60	1	0	16, 20	1	0	10, 12	8
			1	0	16, 20	2	40, 120	10, 12	16
			2	80, 140	16, 20	1	0	10, 12	16
			2	80, 140	16, 20	2	40, 120	10, 12	32
	1.25	30, 55	1	0	20, 30	3	60, 100	8, 10	16
			2	60, 160	20, 30	3	60, 100	8, 10	32
			1	0	20, 30	4	60	8, 10	8
			2	60, 160	20, 30	4	60	8, 10	16
500	0.7	25, 60	1	0	12, 20	1	0	10, 12	8
			1	0	12, 20	2	90, 180	10, 12	16
			2	100, 210	12, 20	1	0	10, 12	16
			2	100, 210	12, 20	2	90, 180	10, 12	32
	1.2	30, 55	1	0	25, 30	3	40, 80	10, 12	16
			2	80, 160	25, 30	3	40, 80	10, 12	32
			1	0	25, 30	6	40	10, 12	8
			2	80, 160	25, 30	6	40	10, 12	16
600	0.8	25, 60	1	0	16, 25	1	0	8, 10	8
			1	0	16, 25	2	80, 120	8, 10	16
			2	100, 180	16, 25	1	0	8, 10	16
			2	100, 180	16, 25	2	80, 120	8, 10	32
	1.1	45, 55	1	0	20, 30	3	60, 100	10, 12	16
			2	80, 200	20, 30	3	60, 100	10, 12	32
			1	0	20, 30	5	60	10, 12	8
			2	80, 200	20, 30	5	60	10, 12	16
Total No. of FE models									720

h = dapped-end depth, a = shear span, d = effective depth, f_c = compressive strength of concrete, φ_h = flexural rebar diameter, φ_v = hanger rebar diameter, N_h = number of flexural layers, S_h = spacing between flexural layers, N_v = number of hanger layers, S_v = spacing between hanger layers.

discriminator. It processes, as an input, a random noise vector and creates output that resembles authentic data. In contrast, the task of the discriminator is to figure out the differences between real and artificially generated samples, assigning probability scores to indicate the likelihood of a sample being real. Throughout the training process, the generator and discriminator engage in an adversarial dynamic, where each component progressively improves in response to the other. As the generator produces increasingly realistic synthetic samples, the discriminator sharpens its ability to differentiate between authentic and generated data. The binary cross-entropy loss function of GAN is denoted by Equation (1), where x is a real sample following the real data distribution Pr, z is a noise vector sampled from the noise distribution Pz, $G(z)$ represents the generated (fake) sample, and D denotes the discriminator function with an output 1 if the given data is (100%) real and 0 if the given data is (100%) fake [43].

$$V_L(D,G) = E_{x\sim Pr}[\log D(x)] + E_{z\sim Pz(z)}[\log(1 - D(G(z)))] \tag{1}$$

The conflicting objectives between the generator and discriminator lead to a competitive training process, which ultimately results in a generator capable of producing realistic-looking data, which is typically the primary objective of using a GAN. CTGAN [51] is an extension of GAN specifically developed for tabular data, which has the ability to generate both continuous and discrete data using conditional sampling and mode-specific normalisation. The fundamental architecture of CTGAN is illustrated in Fig. 7.

The architecture comprises a conditional generator and a discriminator, both implemented using fully connected layers. The generator receives as input a combination of random noise and a conditional vector that specifies the category of the data to be generated. The mode-specific normalisation technique is designed to deal with the multi-modal distribution of continuous variables. For each continuous feature, CTGAN utilises a variational Gaussian mixture model to estimate the number of modes and fit a Gaussian mixture. Each value in the column is then normalised based on the parameters of the corresponding mode it belongs to. The conditional generator and training-by-sampling approach is designed to address the challenge of learning from imbalanced discrete columns. Through conditional sampling during training, CTGAN enables the generator to produce data conditioned on specific discrete values. When generating synthetic data after the training process, CTGAN reverses the transformations and converts the generated samples back to their original scale to ensure that the synthetic data reflects the original feature space.

2.4. CTGAN-based synthetic data generation

The CTGAN-based data augmentation process includes key steps to handle tabular data containing both continuous and categorical variables. In practical applications, the number of hanger and flexural layers (denoted as N_v and N_h) cannot take decimal values and are therefore treated as categorical variables. Similarly, the spacing between hanger (S_v) and flexural (S_h) layers is typically rounded to the nearest 5 mm in design practice. Therefore, these variables are also considered as categorical variables to reflect practical reinforcement detailing. The quality of the synthetic data generated by the generator in CTGAN is highly influenced by the hyperparameters used in the adversarial training process [56]. Therefore, a tuning process was conducted to determine the optimal combination of CTGAN hyperparameters through a grid search approach. The selected key hyperparameters are listed in Table 3. Following 2000 epochs of adversarial training, the generator was able to capture the distribution patterns of the original dataset, as illustrated in Fig. 8. The generator and discriminator losses stabilise after the initial training phase and fluctuate within a narrow range without divergence or mode collapse, indicating a stable adversarial equilibrium. Based on this convergence behaviour, the trained CTGAN model was used to generate synthetic data to enhance the original database.

A total of 680 synthetic data points were generated, aiming to increase the total dataset size to 1400 to improve the performance of the ML model. A comparative analysis of the statistical characteristics between the original and final synthetic data was conducted, as illustrated in Fig. 9.

Fig. 9 demonstrates a comparison of the Kernel Density Estimate (KDE) distributions for the original and synthetic data, where the

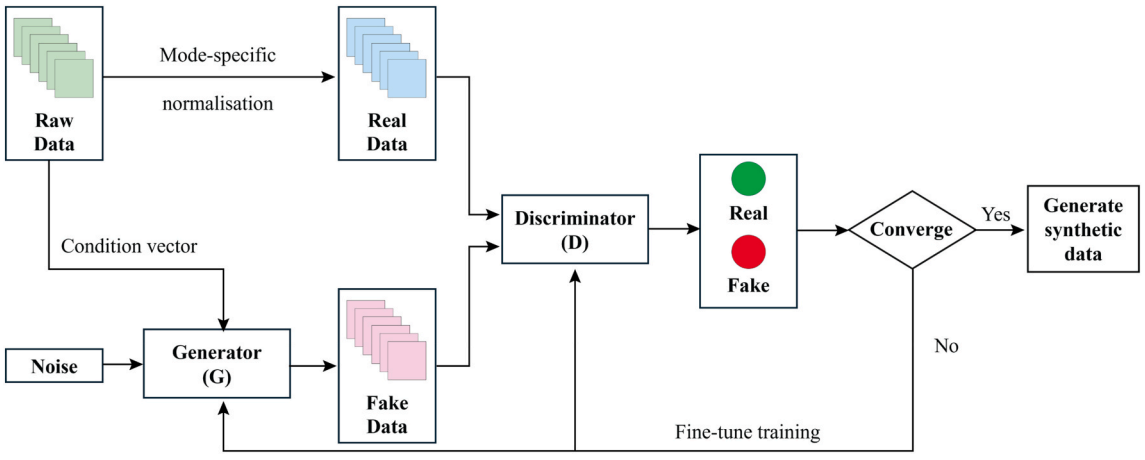


Fig. 7. Architecture of conditional tabular generative adversarial networks (CTGAN).

Table 3
Hyperparameters of the CTGAN.

Hyperparameters	Grid search ranges	Ultimate selection
Size of the random noise vector	32, 64, 128	64
Generator hidden layer size	(32 × 32), (64 × 64), (128 × 128), (256 × 256)	(64 × 64)
Discriminator hidden layer size	(32 × 32), (64 × 64), (128 × 128), (256 × 256)	(64 × 64)
Generator learning rate	1×10^{-5}, 2×10^{-5}, 1×10^{-4}, 2×10^{-4}	2×10^{-5}
Discriminator learning rate	1×10^{-5}, 2×10^{-5}, 1×10^{-4}, 2×10^{-4}	2×10^{-5}
Packed sample size	2, 4, 8	2
Batch size	32, 64, 128	64

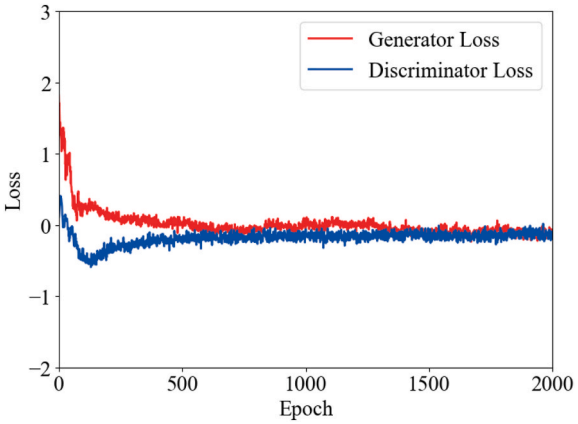


Fig. 8. Generator and discriminator loss curves during CTGAN training.

distribution of the synthetic dataset is closely aligned with that of the original dataset. This indicates that CTGAN has effectively captured the distributional characteristics of the original data. This observation is further supported by the closely aligned values of the mean, standard deviation (Std.), and coefficient of variation (COV) for each parameter. Therefore, the combined dataset containing 1400 data instances was used to train the ML models, while the CTGAN generator was trained using the numerical dataset to capture the overall data distribution and generate realistic synthetic samples. To avoid redundancy and ensure the uniqueness of data points, all generated synthetic samples were screened against the numerical dataset to check for duplicate or overlapping samples with identical input parameter combinations.

3. Development of the surrogate model

The authors selected standalone and hybrid artificial neural network models to predict the capacity of dapped-end connections. This choice is inspired by the fact that hybrid neural network methods have demonstrated high accuracy in various engineering studies compared to classical ML models and individual neural networks [57–59]. In this study, Genetic Algorithm-Artificial Neural Network (GA-ANN) and Particle Swarm Optimisation-Artificial Neural Network (PSO-ANN) were trained, along with an individual Artificial Neural Network (ANN), to compare the predictive capability of hybrid ANN models against individual ANN in estimating the capacity of dapped-end connections. Furthermore, Shapley Additive Explanations (SHAP) were employed to interpret the reasoning behind the predictions of the best-performing model.

3.1. Artificial neural network (ANN)

ANNs are inspired by the function of the human brain and represented through neurons connected in a layered configuration [60]. The grid search method was employed to fine-tune key hyperparameters such as the number of hidden layers, number of neurons per layer, batch size, learning rate, and epochs. An ANN with three hidden layers, each containing 12 neurons was used as the individual ANN. The ReLU activation function was utilised for the hidden layers due to its effectiveness in handling non-linearity. The model was trained using a learning rate of 0.001, 50 epochs and a batch size of 32. To optimise weight updates and improve convergence, the Adam optimiser was used.

3.2. Model optimisation using metaheuristic approach

3.2.1. Genetic Algorithm-Artificial Neural Network (GA-ANN)

To optimise the neural network, a fitness function is defined that builds, trains, and evaluates an ANN using specified

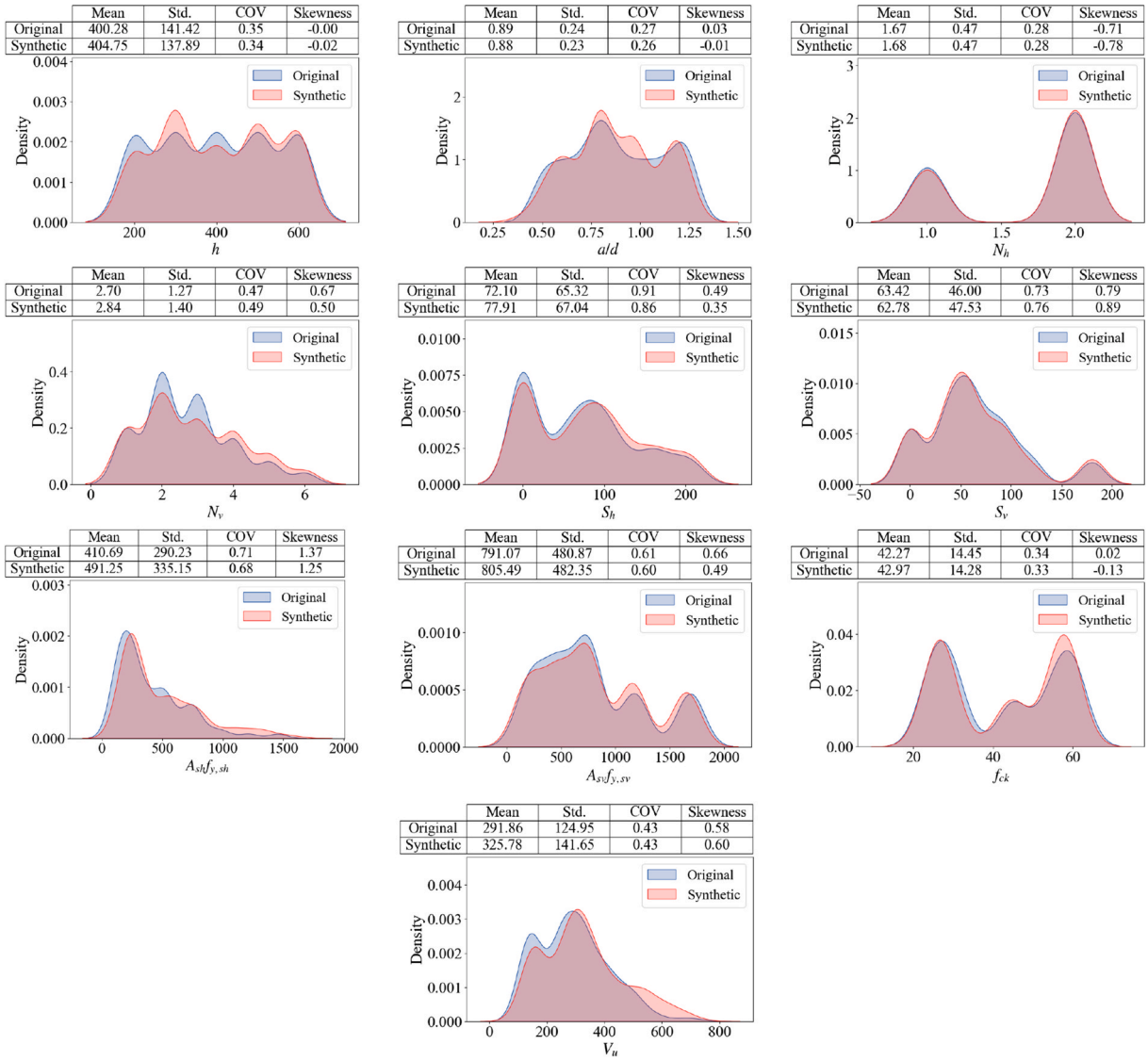


Fig. 9. Comparison of the distributions of parameters in the original and synthetic databases (*Note: The table above each plot presents a comparison of statistical characteristics)*.

hyperparameters. A Genetic Algorithm (GA) begins with a randomly generated population of hyperparameter sets. Each set is evaluated using the fitness function. The GA then applies selection, crossover, and mutation to evolve better solutions over successive generations [61]. Ultimately, the best-performing individual represents the optimal hyperparameters, which are used to configure the final ANN model. A population size of 50, a maximum of 500 iterations, a crossover fraction of 0.8, and 50 generations were used. The GA operates outside the neural network but interacts with it to find the best configuration.

3.2.2. Particle Swarm Optimisation-Artificial Neural Network (PSO-ANN)

Particle Swarm Optimisation employs a population of candidate solutions that navigate the parameter space in search of optimal configurations. Individual particles update their locations by combining personal best discoveries with information shared from the swarm collective, targeting the identification of superior neural network architectures [62]. An evaluation function was developed to assess network performance quality. This modifies the number of neurons in hidden layers according to PSO-generated values, executes the training process, and measures the resulting model's performance. The PSO algorithm was set up with a swarm size of 30, a maximum of 200 iterations, acceleration factors of $c_1 = 1.5$ and $c_2 = 2.5$, and an inertia weight of 0.9. The swarm size represents the number of particles, and the maximum iterations indicate the number of times the particles will try different solutions. For detailed implementation and formulations of the machine learning models, readers are referred to Refs. [63–66].

3.3. Data processing and model optimisation

Data preprocessing is crucial for model training and prediction performances. In this study, several data preprocessing steps were followed to ensure the quality and consistency of the dataset used for machine learning model training. These steps included outlier filtering, data standardisation, data partitioning, and hyperparameter optimisation. Since synthetic data generation may amplify noise or produce unrealistic samples that do not follow the distribution of the original data, an outlier filtering step was applied to the generated synthetic dataset before model training. Similar concerns regarding unrealistic synthetic samples have been noted in previous studies [67,68]. In this study, a synthetic sample was excluded if any of its feature values lay outside the range defined by the mean ± 2 standard deviations, to remove synthetic data points located far from the mean value of the distribution.

As reported in the authors' previous study [69], metaheuristic optimisation techniques such as Genetic Algorithm (GA) and Particle Swarm Optimisation (PSO) performed comparatively better than exhaustive search techniques such as grid search. Therefore, GA and PSO were used for ANN architecture optimisation. For ANN models, as they approximate a continuous non-linear function, data normalisation helps improve the convergence of the training process by ensuring that features are on a comparable scale, facilitating efficient optimisation of the network weights. Data (inputs) standardisation was performed using z-score standardisation (available in Scikit-learn as StandardScaler) so that each feature was rescaled based on its mean and standard deviation, resulting in a transformed distribution with a mean of 0 and a standard deviation of 1.

To evaluate the influence of data augmentation on model performance and generalisation, two separate training cases were considered. In Model 1, the machine learning models were trained using the numerical database containing 720 instances and tested using 64 unseen experimental results collected from the literature. In Model 2, the models were trained using the combined database consisting of 720 numerical instances and 680 synthetic instances and were again tested using the same 64 unseen experimental results. This method was used because the CTGAN was trained using the full numerical database, and the synthetic samples were generated based on the patterns learned from that dataset. Therefore, partitioning the augmented database into training and testing subsets could introduce potential data leakage and uncertainty regarding the independence of the test data in Model 2. Therefore, to ensure a reliable comparison between Model 1 and Model 2, a completely unseen experimental dataset was used for model testing.

3.4. Performance evaluation

Performance assessment plays a critical role in identifying the most effective machine learning algorithm. It confirms that the algorithm has successfully captured underlying data relationships and maintains predictive reliability when applied to previously unencountered samples. Since the synthetic data were generated from numerical samples using CT-GAN, the entire augmented dataset was utilised for model training. To rigorously evaluate generalisation performance an external experimental dataset consisting of 64 unseen test results was exclusively used for model testing. Three metrics were used to assess the performance of ML models (Equations (2)–(4)). These three metrics are Coefficient of Determination (R^2), Normalised Root Mean Squared Error (NRMSE) and Normalised Mean Absolute Error (NMAE). where y denotes the actual data and $\hat{y}$ represents the predicted value. $\bar{y}$ corresponds to the mean value of the actual data and n is the size of the data sample.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{2}$$

$$\text{NRMSE} = \frac{\sqrt{\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}}{\bar{y}} \tag{3}$$

Table 4
Summary of the experimental test results.

Ref.	*h* (mm)	*a/d*	f_c (MPa)	N_h	S_h (mm)	$A_{sh}\,f_{y,sh}$ (kN)	N_v	S_v (mm)	$A_{sv}\,f_{y,sv}$ (kN)	V_u, (kN)	*No. of tests*
[1]	500	0.7	49.6-56.8	1	0	243-1061	2, 3, 4, 6	40, 60, 90	157-1456	245-995	8
[6]	300	0.8	30-41.1	1	0	133-442	1, 3	0, 40	63-293	121-309	9
[17]	325	0.56	33.8	1	0	180	1	0	339	123	1
[24]	300-430	0.61-1.26	27.7-60.6	3	90, 130	720-780	3, 5, 7	16-44	563-896	359-1046	15
[26]	210	0.36, 0.72	37.9-52.9	1	0	78-389	2	25	233-466	136-326	4
[27]	300	0.3, 0.55, 0.6	33.7-69.2	3	90	370-502	4, 5, 6	20	307-654	458-713	9
[28]	279	0.23	35.5	1	0	157	3	50	58.7	189	1
[29]	300	1.2-1.51	32.5-62.9	3	87	367-664	2, 3	25	229-458	234-492	17
Total Results											64

h = dapped-end depth, a = shear span, d = effective depth, f_c = compressive strength of concrete, N_h = number of flexural layers, S_h = spacing between flexural layers, A_{sv} = hanger reinforcement area, N_v = number of hanger layers, S_v = spacing between hanger layers, A_{sv} = hanger reinforcement area, V_u = ultimate capacity of the connection.

$$\mathrm{NMAE} = \frac{\frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i|}{\bar{y}} \tag{4}$$

4. Results and discussion

4.1. Performance of the ML models based on numerical and augmented data

The performance of selected machine learning models was evaluated to determine their predictive capability and generalisation. Since the synthetic data were generated from numerical samples using CTGAN, the entire augmented dataset was utilised for model training. To assess generalisation, model performance was evaluated against an external experimental dataset comprising 64 unseen test results. A summary of the experimental data collected from eight different studies is presented in Table 4.

Table 5 presents the performance indices for training and testing for models trained using the full combined database (1400 instances) and tested against an unseen experimental dataset (64 experimental results). It is observed that the GA-ANN model achieved the best performance in predicting the capacity of dapped-end connections for both training and testing datasets. This indicates that the predictions of the hybrid GA-ANN model generalise well to unseen data. When comparing the R^2 values of the models, the GA-ANN model records the highest R^2 score of 0.973 for training and 0.961 for testing. PSO-ANN also shows good performance in predicting the dapped-end capacity with a training R^2 value of 0.965 and an average testing R^2 value of 0.944. ANN shows a testing R^2 value of 0.901, whereas the training R^2 is 0.886. This indicates that the GA and PSO have optimised the ANN architecture to enhance the performance of the ANN model.

As previous studies have reported that GAN-based data augmentation can improve prediction accuracy and model generalisation [46–50], the benefit of CTGAN augmentation was assessed in this study for predicting the capacity of dapped-end connections. The predictive performance of models trained using the full combined dataset (numerical and synthetic) was compared with that of models trained using the numerical dataset only. In both cases, performance was evaluated using the same unseen experimental dataset. Fig. 10 shows a comparison of predictions for the unseen experimental dataset for each model trained using the combined dataset and the numerical dataset.

Among the models, GA-ANN achieved the highest predictive performance, with an R^2 of 0.961 for the model trained with the combined database and 0.807 for the one trained with the numerical database. PSO-ANN recorded the second-highest accuracy, with R^2 values of 0.944 and 0.706 for the combined and numerical databases, respectively. The standalone ANN model showed the lowest R^2 values in both cases, highlighting that the GA and PSO algorithms effectively optimised the ANN architecture to improve its predictive accuracy and generalisation capability. These results highlight the significant value of using CTGAN-based data augmentation as a solution to data scarcity, enabling the development of more robust and generalisable ML models in structural engineering applications.

4.2. Influence of input parameters on the ultimate capacity of dapped-end connections

After confirming the best model, an investigation was conducted to understand how each parameter influences the dapped-end connection capacity to explore these influences, SHAP values were used as a machine learning explainability method. The SHAP global explanation is shown in Fig. 11. Herein, in a SHAP global explanation, the model-in-whole (entire dataset) is interpreted. In Fig. 11, the x-axis represents the SHAP values, which indicate the contribution of each parameter to the model prediction (capacity of the dapped-end connection). A positive SHAP value signifies that the feature (input parameter) increases the predicted value, while a negative SHAP value means it decreases the prediction. Magnitudes of the SHAP values indicate the degree to which individual variables affect model outputs. Features are arranged vertically in order of decreasing significance, with the most influential parameters positioned at the top of the figure and less impactful variables located toward the bottom. Color-coding represents the numerical range of each input parameter. Higher parameter values are illustrated using red coloration, while lower values appear in blue. For example, the dapped-end depth (h) spans from 190 to 600 within the data collection. Values approaching the upper boundary of 600 display in red, whereas those near the lower threshold of 190 appear in blue.

According to the SHAP analysis, the strength of horizontal (flexural) reinforcement ($A_{sh}f_{y,sh}$) is the most influential parameter governing the capacity of dapped-end connections, followed by the strength of vertical (hanger) reinforcement ($A_{sv}f_{y,sv}$). Experimental

Table 5
Model performance of the model trained using the combined dataset *(Note: training: 1400 numerical and synthetic instances, testing: 64 experimental results)*.

Dataset	Metric	GA-ANN	PSO-ANN	ANN
Training (combined data set 1400)	R^2	0.973	0.965	0.901
	NRMSE	0.081	0.098	0.135
	NMAE	0.051	0.069	0.094
Testing (unseen experimental data 64)	R^2	0.961	0.944	0.886
	NRMSE	0.101	0.120	0.172
	NMAE	0.071	0.081	0.126

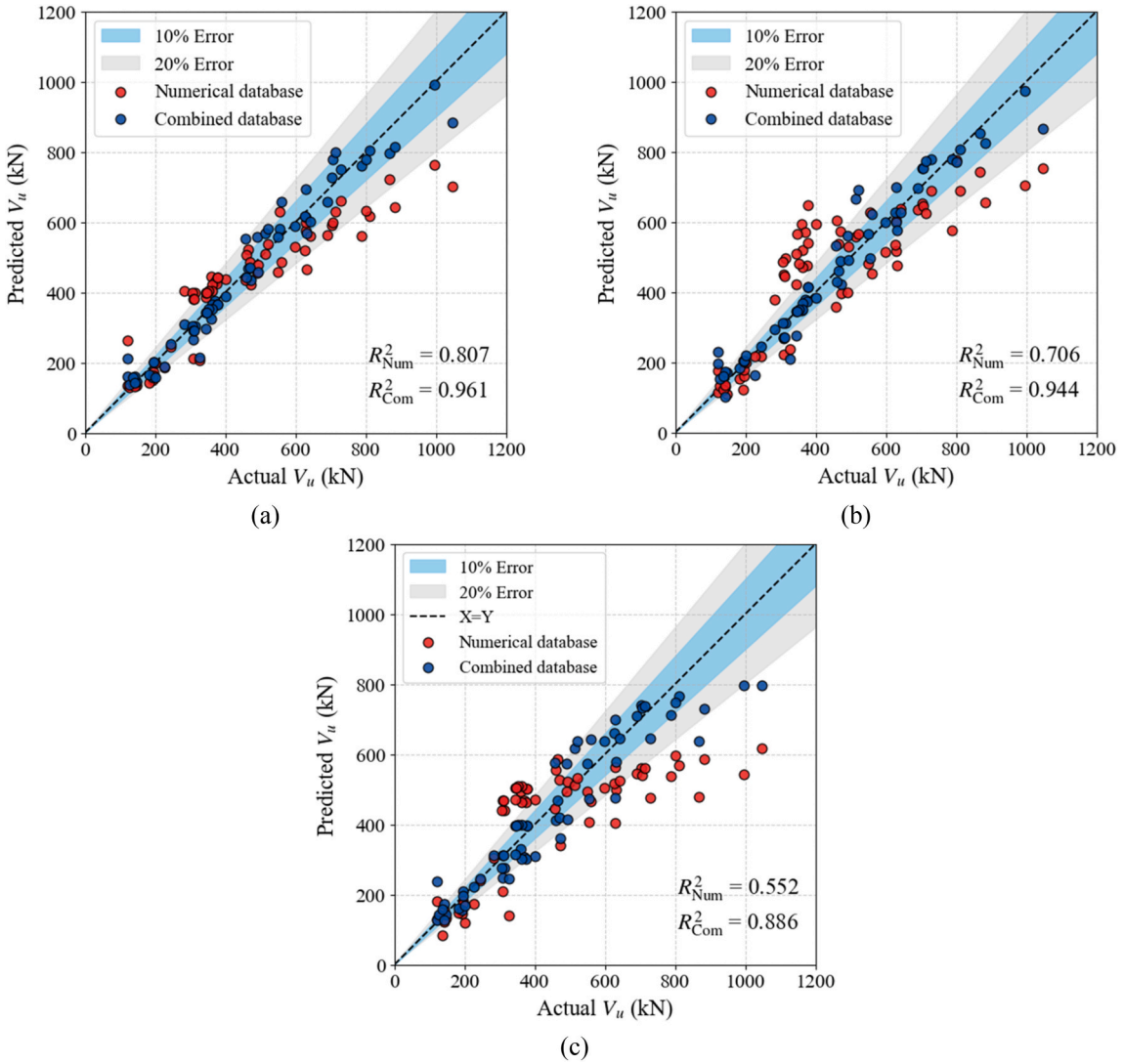


Fig. 10. Comparison of the prediction of the capacity of dapped ends for unseen experimental results: (a) GA-ANN, (b) PSO-ANN and (c) ANN (*Com: refers to the models trained on both numerical & augmented data*).

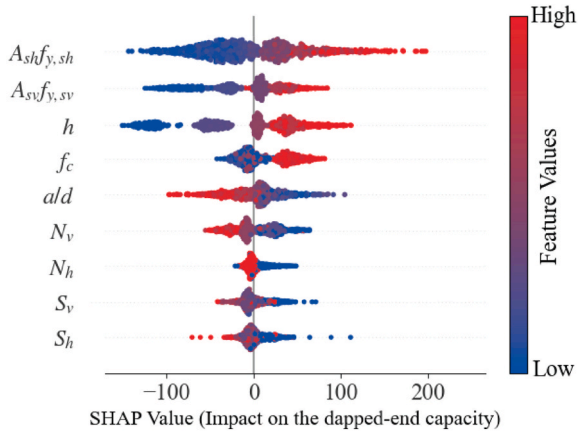


Fig. 11. Influence of input parameters for dapped-end capacity prediction.

evidence has also confirmed the dominance of $A_{sh}f_{y,sh}$ over $A_{sv}f_{y,sv}$, for the same total area of horizontal and vertical reinforcement, connections with more horizontal reinforcement were stronger [70]. Therefore, it indicated that the train GA-ANN provides a reasonable representation of the actual response of dapped-end connections. Furthermore, as the feature values of $A_{sh}f_{y,sh}$ and $A_{sv}f_{y,sv}$ increase, the corresponding SHAP value also increase, which indicates that an increase in both $A_{sh}f_{y,sh}$ and $A_{sv}f_{y,sv}$ enhances the capacity of the dapped-end connections. Similarly, the compressive strength of concrete (f_c) and dapped-end depth (h) follow the same pattern, where an increase in f_c and h enhances the capacity of dapped-end connections. The span-to-effective-depth ratio (a/d) and the spacing between horizontal (S_h) and vertical (S_v) reinforcement layers show a negative relationship with the capacity of dapped-end connections, where an increase in these parameters reduces the capacity of the connection.

To evaluate whether the ML model follows the compatibility of deformations of dapped-end connections, SHAP explanations are compared with the kinematic-based model principles developed by Rajapakse et al. [16]. Which was developed based on first principles, equilibrium, compatibility of deformations, and material constitutive relationships. The model involves an iterative process to determine the depth of the compression zone (x - See Fig. 12 by considering horizontal equilibrium. Once x is obtained, the ultimate capacity of the connection is given by Equation (5).

$$V = \frac{1}{(a + l_{cr}\cos\theta)}\Big[\sum F_{h,i}(h - x - c_{h,i}) + \sum F_{v,i}(l_{cr}\cos\theta - c_{v,i}) + \sum F_{d,i}\cos\beta\sin\theta(l_{cr} - c_{d,i})\sum F_{d,i}\sin\beta\cos\theta(l_{cr} - c_{d,i}) + 0.625F_{c0}(x - d_{cdz}) + F_{CDZ}(x - 0.5d_{cdz}) - H(l_{cr}\sin\theta)\Big] \quad (5)$$

In Equation (5), the forces on the hanger bars and flexural bars are proportional to the capacity of the connection. Therefore, an increase in the strength of the hanger and flexural bars increases the capacity of the dapped-end connection. On the other hand, when the a/d ratio increases, the lever arm to the support vertical reaction increases (see Fig. 12), which leads to a reduction in the dapped-end connection capacity. The inverse relationship of S_v and S_h indicates that larger spacings between layers reduce the capacity of the dapped-end connection. This explanation can be justified using kinematic-based model; when S_v and S_h increase, the lever arm to the corresponding $F_{v,i}$ and $F_{h,i}$ reduces (see Fig. 12). Therefore, an increase in S_v and S_h reduces the capacity of the connection. The SHAP analysis demonstrates that the GA-ANN model effectively captures the underlying relationships between input and output parameters which aligns with the expected behaviour of dapped-end connections.

4.3. Comparative analysis of the neural network predictions and analytical methods on unseen experimental data

The predictions of the GA-ANN model discussed in section 4.1 (trained using combined dataset) were compared with two analytical approaches, which are Strut-and-Tie Model (STM) proposed by Schlaich et al. [3] and the PCI design method [8] to further assess the reliability and accuracy of the ML predictions against analytical methods. As illustrated in Fig. 13, GA-ANN predictions are closely distributed around the ideal value of 1.0, with the smallest Coefficient of Variation (COV) of 14.2% and a mean $V_{u,Act}/V_{u,Pred}$ ratio of 1.00, reflecting more accurate and unbiased prediction. In contrast, both analytical methods underpredict the capacity of dapped-end connections. The PCI method records a mean $V_{u,Act}/V_{u,Pred}$ ratio of 1.42 and the STM method records 1.35, indicating consistent underestimation of actual strength across many data points. Moreover, the prediction scatter of PCI and STM is significantly wider, with COV values of 25.30% and 32.46%, respectively.

These results demonstrate that the GA-ANN model effectively captures the underlying complex interactions among input parameters, leading to more reliable predictions of the capacity of dapped-end connections. While accurate prediction of the capacity of

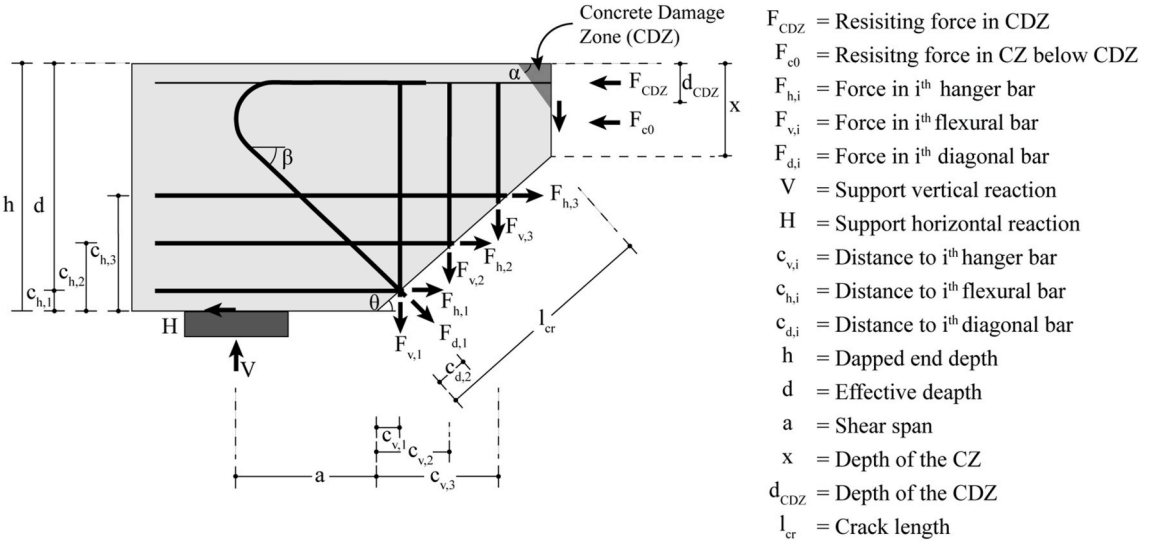


Fig. 12. Forces acting on the free body diagram bounded by the corner crack – kinematic based model [16].

	PCI	STM	GA-ANN
Mean	1.42	1.35	1.00
COV (%)	25.30	32.46	14.20

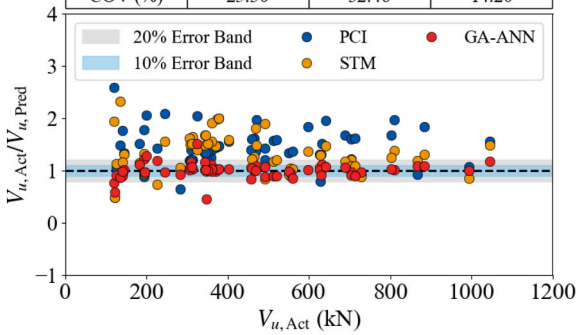


Fig. 13. Comparison of the ML prediction of the capacity of dapped ends for unseen experimental results with analytical methods.

dapped-end connections remains essential, further benefits can be gained from the development of an ML-assisted design model to predict the suitable reinforcement configurations to achieve a user-given dapped-end capacity, which is particularly valuable for practical design and optimisation in civil engineering applications. It is noteworthy that analytical models such as the STM design approach are based on the lower-bound theorem of limit analysis and therefore provide conservative estimates of load-carrying capacity. The GA-ANN model shows closer agreement with experimentally observed capacities, as illustrated in Fig. 13. Since the GA-ANN model is employed in the ML-assisted design framework (Section 4.4), this closer agreement enables the identification of more optimised reinforcement detailing for a user-specified factored design capacity.

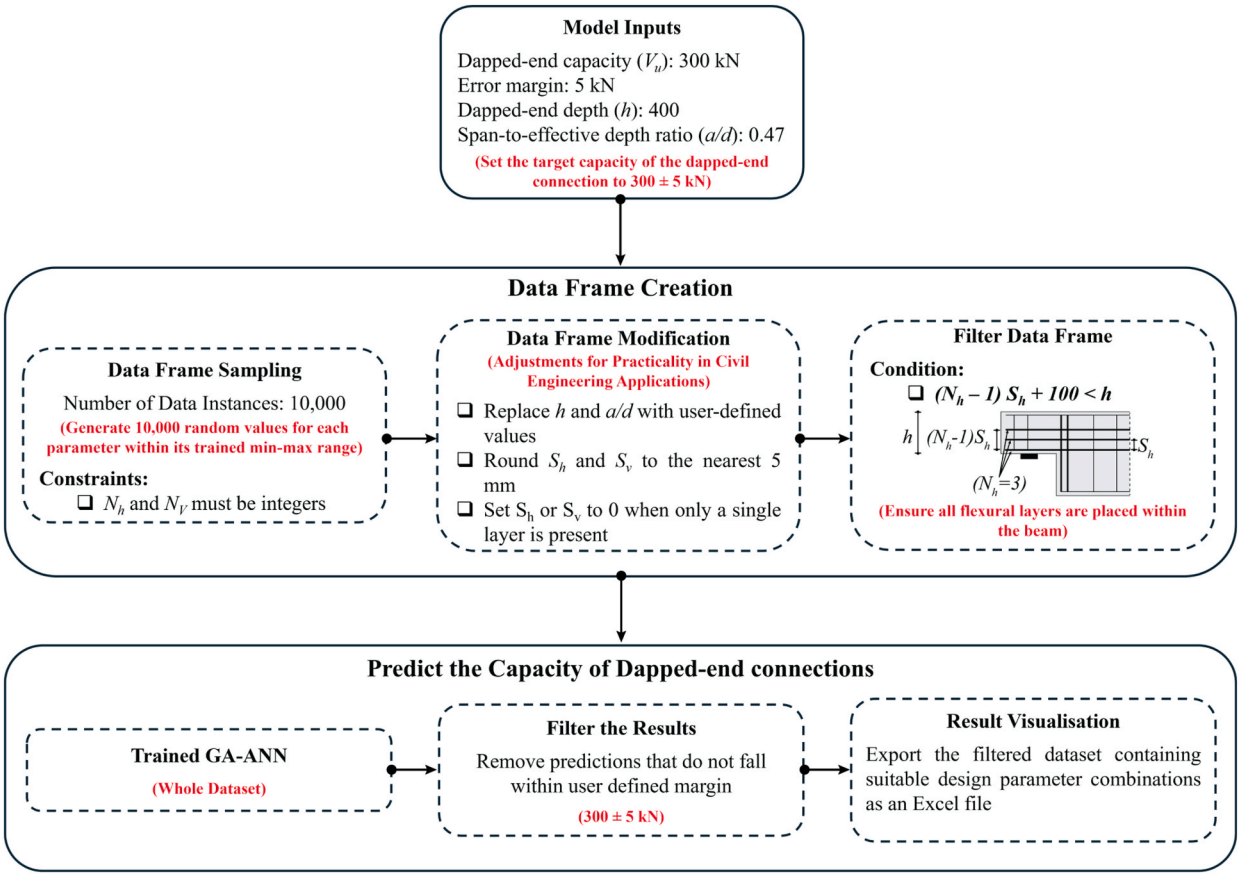


Fig. 14. Methodology for the development of the ML-assisted design model.

4.4. GA-ANN assisted framework for optimisation of the design of dapped end connections

The reverse-engineered workflow of trained ML models is essential in civil engineering applications, as it determines the suitable combinations of design parameters (input parameters) for a given target output (capacity of dapped-end connection) within a specified error margin. In this study, the developed ML-assisted design model takes dapped-end depth (h), span-to-effective-depth ratio (a/d) target load bearing capacity (V_u) of the dapped-end connection and error margin as inputs and then provides suitable design parameters to achieve the design bearing capacity within the error margin.

The methodology followed to develop the ML-assisted design model is illustrated in Fig. 14. The best-performing GA-ANN model was selected as the primary prediction model for this approach. Since the accuracy of the GA-ANN model for the training and testing split has been well proven in this study, the model was retrained using the entire dataset to further enhance its predictive capability for determining design parameters. The first step of this method involved creating a suitable data frame. To achieve this, let P and Q represent the minimum and maximum values of a given input parameter in the dataset used to train the GA-ANN model. A large set of random numbers was generated within the range [P, Q] for each feature using the uniform random sampling technique to create the data frame. This process ensures comprehensive coverage of the multidimensional sample space, as the GA-ANN model has accurately captured the relationships governing the capacity of the dapped-end connections.

In this study, 10,000 random combinations of input parameters were generated. When a large number of random values were generated within the corresponding maximum and minimum ranges of the parameters, the resulting dataset contained decimal values for every parameter. However, decimals for parameters such as number of horizontal (N_h) and vertical (N_v) layers, and the spacing between horizontal (S_h) and vertical (S_v) layers are not meaningful for the real-world application of this ML-assisted model. To ensure practicality, constraints were applied during data frame creation. The number of horizontal (N_h) and vertical (N_v) layers was limited to integer values. For all other parameters, continuous values were randomly generated within their respective [P, Q] range. Then, certain adjustments were made based on the following conditions. If N_h *or* N_v was equal to 1, the corresponding S_h or S_v was set to zero. Furthermore, for better applicability, the values of S_h or S_v were rounded to the nearest 5 mm. Then, the dataset was filtered based on the condition that $(N_h - 1)\, S_h + 100 < h$ to ensure that all flexural layers are placed within the beam. Then. The dapped-end capacity was predicted using the pre-trained GA-ANN model for the parameter combinations in the modified dataset. The algorithm then filtered the dataset by removing combinations where the predicted capacity did not fall within the specified margin.

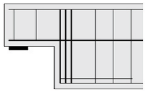
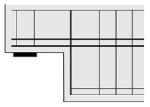
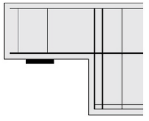
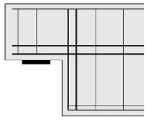
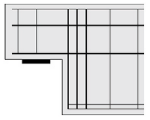
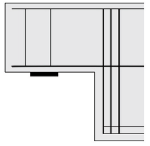
This model can be used as a predictive tool that provides suitable combinations (CM) of design parameters for a user-given h and a/d value, where the predicted capacity of the dapped-end connection falls within the user-specified range. A key advantage of this ML-assisted model is its ability to provide multiple solutions to achieve the target capacity. Table 6 summarises the comparison of ML-assisted model results with VecTor2 FE analysis, where the target capacity was 300 $\pm$ 5 kN. To check the applicability, the results were generated and compared for different h and a/d values. CM4 and CM5 illustrate two different configurations given by the ML-assisted model for a dapped-end depth of 400 mm and an a/d ratio of 0.47 to achieve a capacity of 300 $\pm$ 5 kN. The comparison of FE model results to ML-assisted model predicted capacity ratios shows values of 1.11 and 1.12, respectively. This ability to produce different solutions is particularly valuable in practical design scenarios, as it allows engineers to evaluate alternative solutions and select configurations that best suit construction requirements without compromising structural performance and optimise material usage based on availability and cost. For these six results, the $V_{u,\ FEA}/V_{u,\ Pred}$ ratios record a mean value of 1.07 and COV of 4.10%. These findings indicate that the ML-assisted models offer reliable predictions and can assist in identifying suitable design configurations for dapped-end connections. Interested readers can access the trained GA-ANN model and the developed code for the ML-assisted design model in the GitHub repository [71]. The applicability of the proposed model is limited to dapped-end connections with orthogonal reinforcement layouts and to design parameter combinations that fall within the trained parameter ranges of the model, shown in the distribution plots in Fig. 9. Predictions beyond these ranges are not considered.

5. Conclusion

This paper investigates hybrid deep learning models for predicting the capacity of reinforced concrete dapped-end connections. The numerical database was augmented using a Conditional Tabular Generative Adversarial Network (CTGAN) to enhance data diversity and improve model generalisation. In addition, a deep learning-assisted design framework was developed to estimate design parameters for achieving a target dapped-end capacity within a prescribed error margin, providing multiple feasible design solutions to support optimisation. The following are the key remarks of this study.

- The numerical model was validated using VecTor2 against 24 experimental results, achieving an overall coefficient of variation (COV) of 6.22% between the predicted and measured ultimate capacities. The model also captured key behavioural characteristics, including failure modes, compressive force flow, and crack width development.
- The trained CTGAN generator produced realistic synthetic data consistent with the underlying data distribution, demonstrating its effectiveness in expanding limited datasets. This indicates that CTGAN provides a practical solution for addressing data scarcity in dapped-end connection studies.
- Both hybrid models, GA-ANN and PSO-ANN, outperformed the individual ANN model in predicting the capacity of dapped-end connections, indicating that GA and PSO effectively optimised the hyperparameters, enhancing the predictive capability of the ANN model.
- The SHAP explanation for the best-performing GA-ANN model shows that the influence of each parameter on dapped-end capacity aligns with the underlying principles of the analytical kinematic-based model.

Table 6
Comparison of ML-assisted design model predictions with VecTor2 FEA results for a target ultimate capacity of 300 ± 5 kN.

	Graphical Description	h (mm)	a (mm)	a/d	$A_{sh}\,f_{y,sh}$ (kN)	N_h	S_h (mm)	$A_{sv}\,f_{y,sv}$ (kN)	N_v	S_v (mm)	f_c (MPa)	$V_{u,\ Pred}$ (kN)	$V_{u,\ FEA}$ (kN)	$V_{u,\ FEA}/V_{u,\ Pred}$
CM1		300	230	0.92	480.9 $\varphi = 16$ mm $A_{sh} = 804$ mm^2	1	0	239.7 $\varphi = 10$ mm $A_{sv} = 471$ mm^2	3	40	61.1	304.4	318.5	1.05
CM2		350	235	0.79	480.9 $\varphi = 16$ mm $A_{sh} = 804$ mm^2	2	40	270.53 $\varphi = 12$ mm $A_{sv} = 452$ mm^2	1	0	40.3	302.7	323.4	1.07
CM3		400	305	0.89	563.3 $\varphi = 20$ mm $A_{sh} = 942$ mm^2	1	0	159.8 $\varphi = 10$ mm $A_{sv} = 314$ mm^2	2	60	48.1	299.5	309.2	1.03
CM4		400	160	0.47	364.2 $\varphi = 12$ mm $A_{sh} = 678$ mm^2	2	55	157.1 $\varphi = 8$ mm $A_{sv} = 301$ mm^2	2	55	47.8	300.7	332.9	1.11
CM5		400	160	0.47	443.48 $\varphi = 16$ mm $A_{sh} = 804$ mm^2	2	205	256.3 $\varphi = 10$ mm $A_{sv} = 471$ mm^2	3	65	36.3	302.9	337.8	1.12
CM6		500	320	0.73	242.4 $\varphi = 12$ mm $A_{sh} = 452$ mm^2	1	0	314.2 $\varphi = 8$ mm $A_{sv} = 603$ mm^2	3	60	59.5	304.1	308.3	1.01

h = dapped-end depth, a = shear span, d = effective depth, φ = rebar diameter, A_{sh} = flexural reinforcement area, N_h = number of flexural layers, S_h = spacing between flexural layers.
A_{sv} = hanger reinforcement area, N_v = number of hanger layers, S_v = spacing between hanger layers, f_c = compressive strength of concrete, $V_{u,\ Pred}$ = predicted value by ML-assisted model, $V_{u,\ FEA}$ = estimated through FEA.

- Evaluation against 64 unseen experimental results showed that the GA-ANN model trained using the combined numerical and augmented database achieved superior performance, an R^2 score of 0.961, compared to 0.807 when trained on the numerical database alone. The GA-ANN model also outperformed the STM and PCI analytical methods, achieving a mean $V_{u,Act}/V_{u,Pred}$ ratio of 1.00 and COV of 14.2%, outperforming both analytical approaches.
- A novel ML-assisted design framework was developed to identify design parameters for a target dapped-end capacity within a specified error margin. Multiple feasible design solutions were generated, and six configurations targeting 300 ± 5 kN were validated using finite element analysis in VecTor2, where the $V_{u,\ FEA}/V_{u,\ Pred}$ ratios record a mean value of 1.07 and COV of 4.10%.
- The applicability of the proposed ML-assisted design model is limited to dapped-end connections with orthogonal reinforcement layouts and to design parameter combinations that fall within the trained parameter ranges considered in this study.

CRediT authorship contribution statement

Shashika Dharmawansha: Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Sumudu Herath:** Writing – review & editing, Supervision, Resources, Project administration, Funding acquisition, Conceptualization. **P.L.N. Fernando:** Writing – review & editing, Supervision. **D.P.P. Meddage:** Writing – review & editing. **Chathura Rajapakse:** Writing – review & editing, Supervision, Project administration, Investigation, Formal analysis, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has been supported by the Senate Research Committee (SRC) - Short Term grant (Grant No: SRC/ST/2024/38) of the University of Moratuwa, Sri Lanka.

Data availability

Data will be made available on request.

References

[1] C. Rajapakse, H. Degée, B. Mihaylov, Investigation of shear and flexural failures of dapped-end connections with orthogonal reinforcement, Eng. Struct. 260 (Jun. 2022) 114233, https://doi.org/10.1016/j.engstruct.2022.114233.
[2] J. Mata-Falcón, K.L. Yu, L. Pallarés Rubio, P. Miguel Sosa, Experimental and analytical investigation of the crack behaviour of dapped-end beams, Eng. Struct. 322 (Jan. 2025) 119040, https://doi.org/10.1016/j.engstruct.2024.119040.
[3] J. Schlaich, kurt Schaifer, 'Toward a Consistent Design of Structural Concrete' [Online]. Available: https://www.pci.org/PCI/PCI/Publications/PCI_Journal/Issues/1987/May-June/Toward_a_Consistent_Design_of_Structural_Concrete.aspx.
[4] W.D. Cook, D. Mitchell, Studies of disturbed regions near discontinuities in reinforced concrete members, ACI Struct. J. 85 (2) (1988) 206–216.
[5] J. Mata-Falcón, L. Pallarés, P.F. Miguel, Proposal and experimental validation of simplified strut-and-tie models on dapped-end beams, Eng. Struct. 183 (Mar. 2019) 594–609, https://doi.org/10.1016/j.engstruct.2019.01.010.
[6] J. Mata Falcón, 'Analysis of the Serviceability and Ultimate Behavior of dapped-end Beams (Estudio Del Comportamiento En Servicio Y Rotura De Los Apoyos a Media Madera)', ETH Zurich, 2015 https://doi.org/10.4995/Thesis/10251/53451. Doctoral Thesis.
[7] A. Muttoni, M.F. Ruiz, F. Niketić, M.-R. Backes, Assessment of existing structures based on elastic-plastic stress fields - modelling of critical details and investigation of the in-plane shear transverse bending interaction [Online]. Available: https://infoscience.epfl.ch/server/api/core/bitstreams/8f9faec6-fd93-4a4c-89f8-3db885709be5/content, 2016.
[8] 'PCI design handbook, seventh ed.'. [Online]. Available: https://www.concrete.org/publications/internationalconcreteabstractsportal/m/details/id/51663828.
[9] G. Santarsiero, A. Masi, V. Picciano, Durability of gerber saddles in RC bridges: analyses and applications (Musmeci Bridge, Italy), Infrastructures 6 (2) (Feb. 2021) 2, https://doi.org/10.3390/infrastructures6020025.
[10] T. Nagy-György, G. Sas, A.C. Dăescu, J.A.O. Barros, V. Stoian, Experimental and numerical assessment of the effectiveness of FRP-based strengthening configurations for dapped-end RC beams, Eng. Struct. 44 (Nov. 2012) 291–303, https://doi.org/10.1016/j.engstruct.2012.06.006.
[11] G. Melesse, T. Behailu, H. Kaske Kassa, Finite element analysis of a reinforced concrete dapped-end beam under the effects of impact velocity and dapped-end beam cross-section geometry, Adv. Mater. Sci. Eng. 2023 (1) (2023) 5869552, https://doi.org/10.1155/2023/5869552.
[12] C. Aksoylu, Y.O. Özkılıç, M.H. Arslan, Experimental and numerical investigation of shear strength at dapped end beams having different shear span and recess corner length, Structures 48 (Feb. 2023) 79–90, https://doi.org/10.1016/j.istruc.2022.12.076.
[13] G. Santarsiero, V. Picciano, Post-tension retrofitting of RC dapped-end beams: a numerical investigation, Struct. Concr. 25 (5) (2024) 3246–3264, https://doi.org/10.1002/suco.202300207.
[14] D. D'Angela, G. Magliulo, C. Di Salvatore, E. Cosenza, Computational modeling of reinforced concrete dapped-end beams, Comput. Aided Civ. Infrastruct. Eng. 40 (11) (2025) 1545–1576, https://doi.org/10.1111/mice.13390.
[15] B. Mihaylov, C. Rajapakse, H. Degée, Towards crack-based assessment of dapped-end connections, in: A. Ilki, D. Çavunt, Y.S. Çavunt (Eds.), Building for the Future: Durable, Sustainable, Resilient, Springer Nature Switzerland, Cham, 2023, pp. 352–360, https://doi.org/10.1007/978-3-031-32511-3_38.
[16] C. Rajapakse, H. Degée, B. Mihaylov, Assessment of failure along Re-Entrant corner cracks in existing RC dapped-end connections, Struct. Eng. Int. 31 (2) (Apr. 2021) 216–226, https://doi.org/10.1080/10168664.2021.1878975.
[17] P. Desnerck, J.M. Lees, C.T. Morley, Impact of the reinforcement layout on the load capacity of reinforced concrete half-joints, Eng. Struct. 127 (Nov. 2016) 227–239, https://doi.org/10.1016/j.engstruct.2016.08.061.
[18] P. Desnerck, J.M. Lees, C.T. Morley, The effect of local reinforcing bar reductions and anchorage zone cracking on the load capacity of RC half-joints, Eng. Struct. 152 (Dec. 2017) 865–877, https://doi.org/10.1016/j.engstruct.2017.09.021.

[19] V. Masėnas, A. Meškėnas, J. Valivonis, Analysis of the bearing capacity of reinforced concrete dapped-end beams, Appl. Sci. 13 (9) (Apr. 2023), https://doi.org/10.3390/app13095228.
[20] P. Desnerck, J.M. Lees, C.T. Morley, Strut-and-tie models for deteriorated reinforced concrete half-joints, Eng. Struct. 161 (Apr. 2018) 41–54, https://doi.org/10.1016/j.engstruct.2018.01.013.
[21] Y.O. Özkılıç, C. Aksoylu, M.H. Arslan, Experimental and numerical investigations of steel fiber reinforced concrete dapped-end purlins, J. Build. Eng. 36 (Apr. 2021) 102119, https://doi.org/10.1016/j.jobe.2020.102119.
[22] S.E.-D.M.F. Taher, Strengthening of critically designed girders with dapped ends, Proc. Inst. Civ. Eng. - Struct. Build. 158 (2) (Apr. 2005) 141–152, https://doi.org/10.1680/stbu.2005.158.2.141.
[23] J.Y. Moreno-Martínez, R. Meli, Experimental study on the structural behavior of concrete dapped-end beams, Eng. Struct. 75 (Sep. 2014) 152–163, https://doi.org/10.1016/j.engstruct.2014.05.051.
[24] W.-Y. Lu, I.-J. Lin, H.-W. Yu, Behaviour of reinforced concrete dapped-end beams, Mag. Concr. Res. 64 (9) (Sep. 2012) 793–805, https://doi.org/10.1680/macr.11.00116.
[25] M. Aswin, A. Al-Fakih, Z.I. Syed, M.S. Liew, Influence of different dapped-end reinforcement configurations on structural behavior of RC dapped-end beam, Buildings 13 (1) (Jan. 2023), https://doi.org/10.3390/buildings13010116.
[26] L.A. Clark, P. Thorogood, Serviceability behaviour of reinforced concrete half joints, Struct. Eng. 66 (18) (1988) 295–302.
[27] W. Lu, I. Lin, S. Hwang, Y. Lin, Shear strength of high-strength concrete dapped-end beams, J. Chin. Inst. Eng. 26 (5) (Jul. 2003) 671–680, https://doi.org/10.1080/02533839.2003.9670820.
[28] S. Ahmad, A. Elahi, J. Hafeez, M. Fawad, Z. Ahsan, Evaluation of the shear strength of dapped ended beam, Life Sci. J. 10 (Jul. 2013) 1038–1044.
[29] W.-Y. Lu, T.-C. Chen, I.-J. Lin, Shear strength of reinforced concrete dapped-end beams with shear span-to-depth ratios greater than Unity, J. Mar. Sci. Technol. 23 (4) (Aug. 2015), https://doi.org/10.6119/JMST-015-0511-1.
[30] A. Atta, M. Taman, Innovative method for strengthening dapped-end beams using an external prestressing technique, Mater. Struct. 49 (8) (Aug. 2016) 3005–3019, https://doi.org/10.1617/s11527-015-0701-8.
[31] H. Naser, Q. Shakir, Experimental Study of the Behavior of Reinforced Concrete Beams with Composite Dapped End Under Effect of Static and Repeated Loads, Mar. 2019, https://doi.org/10.30560/ijas.v2n1pXX.
[32] G. Menichini, F. Gusella, S.G. Morano, M. Orlando, RC dapped-end beams with various reinforcement layouts: an experimental investigation, Eng. Struct. 322 (Jan. 2025) 119043, https://doi.org/10.1016/j.engstruct.2024.119043.
[33] C. Rajapakse, H. Degee, B. Mihaylov, Experimental investigation of dapped ends with diagonal reinforcement [Online]. Available: https://www.concrete.org/publications/internationalconcreteabstractsportal.aspx?m=details&id=51740710.
[34] D. D'Angela, C. Di Salvatore, M. Acanfora, E. Cosenza, G. Magliulo, Dapped-End beams: experimental tests and capacity models in the literature, in: M.A. Aiello, A. Bilotta (Eds.), Proceedings of Italian Concrete Conference 2022, Springer Nature Switzerland, Cham, 2024, pp. 302–312, https://doi.org/10.1007/978-3-031-43102-9_24.
[35] S. Dharmawansha, N. Silva, N. Kumarasinghe, S. Herath, P.L.N. Fernando, C. Rajapakse, Experimental and numerical investigation of bond conditions in reinforced concrete half-joints, Eng. Struct. 355 (2026) 122447, https://doi.org/10.1016/j.engstruct.2026.122447.
[36] M. Mohtasham Moein, et al., Predictive models for concrete properties using machine learning and deep learning approaches: a review, J. Build. Eng. 63 (Jan. 2023) 105444, https://doi.org/10.1016/j.jobe.2022.105444.
[37] R. Kumar, et al., Machine and deep learning methods for concrete strength prediction: a bibliometric and content analysis review of research trends and future directions, Appl. Soft Comput. 164 (2024) 111956–Oct, https://doi.org/10.1016/j.asoc.2024.111956.
[38] P. Padmapoorani, S. Senthilkumar, R. Mohanraj, Machine learning techniques for structural health monitoring of concrete structures: a systematic review, Iran. J. Sci. Technol. Trans. Civ. Eng. 47 (4) (Aug. 2023) 1919–1931, https://doi.org/10.1007/s40996-023-01054-5.
[39] T. Ali, et al., A systematic literature review of AI-based prediction methods for self-compacting, geopolymer, and other eco-friendly concrete types: advancing sustainable concrete, Constr. Build. Mater. 440 (Aug. 2024) 137370, https://doi.org/10.1016/j.conbuildmat.2024.137370.
[40] N. Jayaweera, A. Abeyrathna, S. Herath, I.U. Ekanayake, D.P.P. Meddage, Code-compliant optimal design of reinforced concrete slab beam systems for low to mid-rise buildings using heterogeneous graph neural networks with attention layers, J. Build. Eng. 117 (Jan. 2026) 114599, https://doi.org/10.1016/j.jobe.2025.114599.
[41] V. Picciano, G. Santarsiero, A. Masi, A. Digrisolo, Structural analysis of dapped-end beams through machine learning techniques, in: Bridge Maintenance, Safety, Management, Digitalization and Sustainability, CRC Press, 2024.
[42] P. Chai, L. Hou, G. Zhang, Q. Tushar, Y. Zou, Generative adversarial networks in construction applications, Autom. Constr. 159 (Mar. 2024) 105265, https://doi.org/10.1016/j.autcon.2024.105265.
[43] I.J. Goodfellow, et al., Generative adversarial nets, in: Advances in Neural Information Processing Systems, Curran Associates, Inc., 2014 [Online]. Available: https://proceedings.neurips.cc/paper_files/paper/2014/hash/f033ed80deb0234979a61f95710dbe25-Abstract.html.
[44] A. Ba Ragaa, F. Al-Neshawy, M. Noureldin, AI-based framework for concrete durability assessment using generative adversarial networks and bayesian neural networks, Constr. Build. Mater. 471 (Apr. 2025) 140722, https://doi.org/10.1016/j.conbuildmat.2025.140722.
[45] N. Chen, et al., Virtual mix design: prediction of compressive strength of concrete with industrial wastes using deep data augmentation, Constr. Build. Mater. 323 (Mar. 2022) 126580, https://doi.org/10.1016/j.conbuildmat.2022.126580.
[46] A. Marani, M.L. Nehdi, Predicting shear strength of FRP-reinforced concrete beams using novel synthetic data driven deep learning, Eng. Struct. 257 (Apr. 2022) 114083, https://doi.org/10.1016/j.engstruct.2022.114083.
[47] S. Wang, H. Yang, T. Stratford, J. He, B. Li, J. Su, Evaluating the effect of curing conditions on the glass transition of the structural adhesive using conditional tabular generative adversarial networks, Eng. Appl. Artif. Intell. 130 (Apr. 2024) 107796, https://doi.org/10.1016/j.engappai.2023.107796.
[48] Y. Wang, et al., Multi-objective optimization of engineered cementitious composite based on machine learning and generative adversarial network, J. Build. Eng. 96 (Nov. 2024) 110471, https://doi.org/10.1016/j.jobe.2024.110471.
[49] S. Zhang, et al., Generative adversarial network-enhanced machine learning models for high-precision prediction of rectangular concrete-filled steel tube strength, Eng. Struct. 339 (Sep. 2025) 120514, https://doi.org/10.1016/j.engstruct.2025.120514.
[50] X. Shi, et al., Improving hydraulic conductivity prediction of bentonite using machine learning with generative adversarial network-based data augmentation, Constr. Build. Mater. 462 (Feb. 2025) 139962, https://doi.org/10.1016/j.conbuildmat.2025.139962.
[51] L. Xu, M. Skoularidou, A. Cuesta-Infante, K. Veeramachaneni, Modeling Tabular data using Conditional GAN, Adv. Neural Inf. Process. Syst. 32 (2019).
[52] S.M. Lundberg, S.-I. Lee, A unified approach to interpreting model predictions, in: Advances in Neural Information Processing Systems, Curran Associates, Inc., 2017 [Online]. Available: https://proceedings.neurips.cc/paper/2017/hash/8a20a8621978632d76c43dfd28b67767-Abstract.html.
[53] P. S. Wong, F. J. Vecchio, and H. Trommels, 'VecTor2 and Formworks User's Manual – second ed.'.
[54] B. Mihaylov, C. Rajapakse, P.-H. Berger, Effect of steel fibers on the ultimate flexural behavior of dapped-end connections, Eng. Struct. 259 (May 2022) 114147, https://doi.org/10.1016/j.engstruct.2022.114147.
[55] F.J. Vecchio, Disturbed stress field model for reinforced concrete: formulation, J. Struct. Eng. 126 (9) (Sep. 2000) 1070–1077, https://doi.org/10.1061/(ASCE)0733-9445(2000)126:9(1070).
[56] A. Marani, L. Zhang, M.L. Nehdi, Design of concrete incorporating microencapsulated phase change materials for clean energy: a ternary machine learning approach based on generative adversarial networks, Eng. Appl. Artif. Intell. 118 (Feb. 2023) 105652, https://doi.org/10.1016/j.engappai.2022.105652.
[57] T. Rajaee, S. Khani, M. Ravansalar, Artificial intelligence-based single and hybrid models for prediction of water quality in rivers: a review, Chemom. Intell. Lab. Syst. 200 (May 2020) 103978, https://doi.org/10.1016/j.chemolab.2020.103978.
[58] J. Sun, L. Fang, J. Han, Optimization of concrete hollow brick using hybrid genetic algorithm combining with artificial neural networks, Int. J. Heat Mass Transf. 53 (23) (Nov. 2010) 5509–5518, https://doi.org/10.1016/j.ijheatmasstransfer.2010.07.006.

[59] A. Behnood, E.M. Golafshani, Predicting the compressive strength of silica fume concrete using hybrid artificial neural network with multi-objective grey wolves, J. Clean. Prod. 202 (Nov. 2018) 54–64, https://doi.org/10.1016/j.jclepro.2018.08.065.
[60] O.A. Montesinos López, A. Montesinos López, J. Crossa, Fundamentals of artificial neural networks and deep learning, in: O.A. Montesinos López, A. Montesinos López, J. Crossa (Eds.), Multivariate Statistical Machine Learning Methods for Genomic Prediction, Springer International Publishing, Cham, 2022, pp. 379–425, https://doi.org/10.1007/978-3-030-89010-0_10.
[61] L.T. Le, H. Nguyen, J. Dou, J. Zhou, A comparative Study of PSO-ANN, GA-ANN, ICA-ANN, and ABC-ANN in estimating the heating load of buildings' energy efficiency for Smart City planning, Appl. Sci. 9 (13) (Jan. 2019) 13, https://doi.org/10.3390/app9132630.
[62] E.T. Mohamad, D. Jahed Armaghani, E. Momeni, S.V. Alavi Nezhad Khalil Abad, Prediction of the unconfined compressive strength of soft rocks: a PSO-based ANN approach, Bull. Eng. Geol. Environ. 74 (3) (Aug. 2015) 745–757, https://doi.org/10.1007/s10064-014-0638-0.
[63] E.H. Houssein, A.G. Gad, K. Hussain, P.N. Suganthan, Major advances in particle swarm optimization: theory, analysis, and application, Swarm Evol. Comput. 63 (Jun. 2021) 100868, https://doi.org/10.1016/j.swevo.2021.100868.
[64] M. Imran, R. Hashim, N.E.A. Khalid, An overview of particle swarm optimization variants, Procedia Eng. 53 (Jan. 2013) 491–496, https://doi.org/10.1016/j.proeng.2013.02.063.
[65] A.G. Gad, Particle swarm optimization Algorithm and its applications: a systematic review, Arch. Comput. Methods Eng. 29 (5) (Aug. 2022) 2531–2561, https://doi.org/10.1007/s11831-021-09694-4.
[66] S. Katoch, S.S. Chauhan, V. Kumar, A review on genetic algorithm: past, present, and future, Multimed. Tools Appl. 80 (5) (Feb. 2021) 8091–8126, https://doi.org/10.1007/s11042-020-10139-6.
[67] M.U. Safder, S.S. Naveed, K. Khurshid, A. Salman, I.F. Nizami, Optimizing imbalanced learning with genetic algorithm, Sci. Rep. 15 (1) (Oct. 2025) 34857, https://doi.org/10.1038/s41598-025-09424-x.
[68] S. Hong, S. An, J.-J. Jeon, Improving SMOTE via fusing conditional VAE for data-adaptive noise filtering, Appl. Intell. 55 (12) (Jul. 2025) 841, https://doi.org/10.1007/s10489-025-06692-y.
[69] H. Sajindra, S. Dharmawansha, H. Wijesundara, S. Herath, U. Rathnayake, D.P.P. Meddage, Hybrid neural network methods to model the external wind pressure on a low-rise flat-roofed building in an irregularly shaped urban environment, Structures 79 (Sep. 2025) 109529, https://doi.org/10.1016/j.istruc.2025.109529.
[70] C. Rajapakse, B. Mihaylov, H. Degée, Behaviour and modelling of reinforced concrete dapped-end connections [Online]. Available: https://www.fsa.uliege.be/cms/c_9825848/en/behaviour-and-modelling-of-reinforced-concrete-dapped-end-connections.
[71] S. Dharmawansha, ML-assisted design model for dapped-end connections, GitHub (2024) [Online]. Available: https://github.com/Shashika-td/ML-assisted-design-model-for-dapped-end-connections-.git.