J. Math. Computer Sci., 40 (2026), 310–329

Online: ISSN 2008-949X

Journal of Mathematics and Computer Science

Journal Homepage: www.isr-publications.com/jmcs

A hybrid technique of deep learning neural networks with finite difference method for higher order fractional Volterra-Fredholm integro-differential equations with φ-Caputo operator

Kawthar Alsa'di[a,*], Nik Mohd Asri B Nik Long[a,b,*]

[a]Department of Mathematics and Statistics, Faculty of Science, Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia.

[b]Laboratory of Computational Sciences and Mathematical Physics, Institute for Mathematical Research, Universiti Putra Malaysia, Serdang 43400, Selangor, Malaysia.

Abstract

This paper deals with the theoretical and numerical aspects for higher order Volterra-Fredholm fractional integro-differential equations (VF-IDEs) under φ-Caputo operator. Using Liptchiz conditions, Krasnoselskii's fixed point theorem, and Gronwall inequality with respect to the function φ, existence and uniqueness of the solution are investigated. The stability of the solution is analyzed through the continuity of the parameters. Moreover, a new hybrid technique which is the combination of deep learning artificial neural network and finite difference method (FDL-ANN) is developed to approximate the solution of higher order VF-IDEs. This technique uses the Adaptive Moment Estimation Method (Adam) as an optimization algorithm with feed-forward deep learning to minimize the error function and training the model using five layers with different activation functions. The numerical analysis for the error bound and the computation complexity are provided for FDL-ANN. The numerical examples demonstrated the efficiency of the proposed method in solving the complicated higher order fractional problems of linear and non-linear terms.

Keywords: Fractional integro-differential equation, φ-Caputo operator, artificial neural network, Adam optimization, deep learning, feed-forward networks.

2020 MSC: 26A33, 45J05, 65L60, 65M06, 68T07.

1. Introduction

Fractional integro-differential equations (FIDEs) have received a great attention to represent some nonlinear phenomenons in variety of fields. Due to the memory effect property of fractional calculus, various types of fractional operators were constructed and used in modeling some real-life problems. For example, Caputo derivative models classical memory with singular kernels (power-law decay), Atangana-Baleanu derivative involves non-singular and non-local kernels, better describing non-singular memory phenomena, Hilfer derivative interpolates between Riemann-Liouville and Caputo derivatives which is useful for systems with mixed memory effects. Combining such of these operators with a strictly monotone increasing function (called φ) are considered as a generalization for many fractional operators such

*Corresponding author

Email addresses: kawther.alsaadi@yahoo.com (Kawthar Alsa'di), nmasri@upm.edu.my (Nik Mohd Asri B Nik Long)

doi: 10.22436/jmcs.040.03.02

Received: 2024-12-29 *Revised:* 2025-03-05 *Accepted:* 2025-05-22

as Caputo fractional operator, Caputo-Hadamard operator, exponential Caputo derivative and offer the flexibility in adapting to specific physical systems. For this reason, many researchers utilized this operator in studying the fractional differential equations (FDEs), for example, φ-Hilfer infinite delay FDEs [1], φ-Caputo unified Maxwell model with a variable kernel [18], φ-Caputo fractional generalized water transport equation [8], φ-Caputo FDEs [23], and φ-FDEs [7].

FIDEs have some complexity in studying their nature, theoretically and numerically especially with respect to another functions. Dhayal and Zhu [11] used a semigroup theory and a fixed-point technique to study the existence and uniqueness of mild solutions for φ-Hilfer impulses fractional integro-differential equations (FIDEs), and proved the stability criteria using the Grönwall inequality and investigated the controllability of the system using a piecewise control function. Zhou et al. [28] used Banach contraction mapping principle to investigate the existence and uniqueness of solution for the nonlinear φ-Hilfer FIDEs on an unbounded domain. Recently, Heydari and Razzaghi [14] used the integral operational matrix of Chebyshev cardinal wavelets with block-pulse functions to solve the φ-Caputo nonlinear FIDEs. Li et al. [19] used the finite difference methods for solving a φ-Caputo nonlinear FDEs in different aspect.

Recently, Artificial Neural Networks (ANN) and deep learning, which consist of more than two hidden layers, are used to solve different type of mathematical models of ordinary differential equations, partial differential equation, and integro-differential equations. For example, Fang et al. [12] used a deep neural networks for solving fractional partial differential equations, Mukdasai et al. [21] designed a supervised neural networks solution for fractional order Leptospirosis biological model, also authors in [13] used ANN to provide different variables for forecasting the cumulative deaths caused by COVID-19 in the USA. Admon et al. [3] used deep feed-forward neural network with vectorized algorithm to solve fractional differential equations, Zúñiga-Aguilar et al. [29] used the ANN with one hidden layer to approximate the solution of fractional differential equations. Also, a linear fractional Volterra type integro-differential equations solved by Allahviranloo et al. [4] using a multi-layered feed-forward neural network, and for two-dimensional equation by Saneifard et al. [24] using ANN. Jafarian and Nia [15] used a combination of neural networks and power series method to solve a fractional Fredholm type integro-differential equation.

Motivated from all of the above mentioned works, we consider a new higher order Volterra-Fredholm via φ-Caputo derivative (φ-Caputo VF-IDEs) with initial boundary conditions:

$$\begin{cases} {}^{\mathcal{C}}\mathfrak{D}^{\theta,\varphi}_{c_0}\chi(u) = \mathfrak{O}\left(u,\chi(u)\right) + \int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv + \int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv, \\ \chi^{(k)}_{\varphi}(c_0) = \rho_k, \ (u \in J = [c_0, C], c_0 \geqslant 0, \ k = 0,1,2,\ldots,n-1), \end{cases} \tag{1.1}$$

where ${}^{\mathcal{C}}\mathfrak{D}^{\theta,\varphi}_{c_0}$ is φ-Caputo fractional derivative of order $\theta \in (n-1,n], n \in \mathbb{N}$. Furthermore, the unknown function $\chi \in C^{n-1}([c_0,C])$ such that the operator ${}^{\mathcal{C}}\mathfrak{D}^{\theta,\varphi}_{c_0}$ exists and ${}^{\mathcal{C}}\mathfrak{D}^{\theta,\varphi}_{c_0}\chi \in C([c_0,C])$, $\rho_k \in \mathbb{R}$, $\mathfrak{O} : J \times \mathbb{R} \to \mathbb{R}$ is a continuous function, $\mathfrak{H}_1, \mathfrak{H}_2$ are linear or nonlinear continuous functions, with $\alpha_1 = \max_{u,v\in J} \int_{c_0}^{C} |\mathfrak{K}_1(u,v)|\,dv$, $\alpha_2 = \max_{u,v\in J} \int_{c_0}^{u} |\mathfrak{K}_2(u,v)|\,dv$, $\mathfrak{K}_1, \mathfrak{K}_2 \in (J \times J, \mathbb{R})$, $\chi^{(k)}_{\varphi}(u) = \left(\frac{1}{\varphi'(u)}\frac{d}{du}\right)^k \chi(u)$, and the increasing function $\varphi'(u) \in C^n([a,b])$ is such that $\varphi'(u) \neq 0$, $\forall\, u \in [c_0, C]$, where $C^{n-1}([c_0,C])$ is the Banach space containing all the $(n-1)$-th-continuously differentiable functions χ such that the norm is defined as $\|\chi\| = \sup_{u\in J} |\chi(u)|$.

To the best of our knowledge, the theoretical and numerical studies for Eq. (1.1) are still not well investigated with respect to another function under higher fractional order. Our work addresses these two gaps and using the FDL-ANN the proposed Eq. (1.1) will be solved.

The main contributions of this study are summarized as follows.

1. The existence and uniqueness of the solution for nonlinear higher order φ-Caputo FIDEs are proved using the Krasnoselskii's fixed point theorem, and the Gronwall inequality with respect to the φ-function under Liptchiz conditions.
2. The stability of solution for the higher order φ-Caputo FIDEs is investigated under the continuity depends on the parameters.

3. A new hybrid technique using FDL-ANN is developed to solve a non-linear higher order φ-Caputo FIDEs.

This paper is organized as follows. In Section 2, we recall some background of φ-Caputo fractional operator in fractional calculus. In Section 3, we investigate the existence, uniqueness, and stability of the solution. Section 4 is dedicated to the construction of the approximate solution using the deep learning artificial neural network. In Section 6, we present some numerical examples to demonstrate the practicality of the developed theorems. Finally, in Section 7, we summarize our work in the conclusion.

2. Preliminaries

In this section, we recall some concepts and theorems related to φ-Caputo fractional operator that are used throughout this paper.

Definition 2.1 ([5, 6]). Let $\theta > 0, \chi$ an integrable function defined on $[c_0, C]$ and $\varphi \in C^n[c_0, C]$ an increasing differentiable function such that $\varphi'(u) \neq 0$ for all $u \in [c_0, C]$. The left-sided φ-Riemann-Liouville fractional of order θ of a function χ is given by

$$I_{c_0^+}^{\theta,\varphi}\chi(u) = \frac{1}{\Gamma(\theta)}\int_{c_0}^{u} \varphi'(s)(\varphi(u)-\varphi(s))^{\theta-1}\chi(s)ds,$$

where $\Gamma(\cdot)$ is a gamma funcion.

Definition 2.2 ([5, 6]). Let $n-1 < \theta < n, \chi : [c_0, C] \to \mathbb{R}$ be an integrable function and φ is defined as in Definition 2.1. The left-sided φ-Riemann-Liouville fractional derivative of order θ of a function χ is defined by

$${}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = \left[\frac{1}{\varphi'(u)}\frac{d}{du}\right]^n I_{c_0^+}^{n-\theta,\varphi}\chi(u),$$

where $n = [\theta]+1$ and $[\theta]$ denotes the integer part of the real number θ.

Definition 2.3 ([5, 6]). Let $n-1 < \theta < n, \chi \in C^{n-1}[c_0, C]$ and φ be defined as in Definition 2.1. The left-sided φ-Caputo fractional derivative of function χ of order θ is determined as

$${}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = {}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\left[\chi(u) - \sum_{k=0}^{n-1}\frac{\chi_\varphi^{[k]}(c_0)}{k!}(\varphi(u)-\varphi(c_0))^k\right],$$

where $\chi_\varphi^{[k]}(u) = \left[\frac{1}{\varphi'(u)}\frac{d}{du}\right]^k\chi(u)$ and $n = [\theta]+1$ for $\theta \notin \mathbb{N}, n = \theta$ for $\theta \in \mathbb{N}$. Further, if $\chi \in C^n[c_0, C]$ and $\theta \notin \mathbb{N}$, then

$${}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = I_{c_0^+}^{n-\theta,\varphi}\left[\frac{1}{\varphi'(u)}\frac{d}{du}\right]^n\chi(u) = \frac{1}{\Gamma(n-\theta)}\int_{c_0}^{u}\varphi'(s)(\varphi(u)-\varphi(s))^{n-\theta-1}\chi_\varphi^{[n]}(s)ds. \tag{2.1}$$

Thus, if $\theta = n \in \mathbb{N}$, one has ${}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = \chi_\varphi^{[n]}(u)$.

Lemma 2.4 ([5, 6]). *Let $\theta > 0$ and $\chi : [c_0, C] \to \mathbb{R}$. The following holds.*

1. *If $\chi \in C[c_0, C]$, then ${}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}I_{c_0^+}^{\theta,\varphi}\chi(u) = \chi(u)$.*
2. *If $\chi \in C^{n-1}[c_0, C]$, then $I_{c_0^+}^{\theta,\varphi}{}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = \chi(u) - \sum_{k=0}^{n-1}c_k[\varphi(u)-\varphi(c_0)]^k$, where $c_k = \frac{\chi_\varphi^{[k]}(c_0)}{k!}$.*

Lemma 2.5 ([16]). *Let $\theta, \gamma > 0$ and $\chi : [c_0, C] \to \mathbb{R}$. Then*

1. $I_{c_0^+}^{\theta,\varphi}[\varphi(u)-\varphi(c_0)]^{\gamma-1} = \frac{\Gamma(\gamma)}{\Gamma(\theta+\gamma)}[\varphi(u)-\varphi(c_0)]^{\theta+\gamma-1}$;

2. ${}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0}[\varphi(u)-\varphi(c_0)]^{\gamma-1} = \frac{\Gamma(\gamma)}{\Gamma(\gamma-\theta)}[\varphi(u)-\varphi(c_0)]^{\gamma-\theta-1}$;
3. ${}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0}[\varphi(u)-\varphi(a)]^{k} = 0, \forall k \in \{0,1,\ldots,n-1\}, n \in \mathbb{N}$;
4. $I^{\theta,\varphi}_{c_0^+} I^{\gamma,\varphi}_{c_0^+}\chi(u) = I^{\theta+\gamma,\varphi}_{c_0^+}\chi(u)$.

Proof. See [16]. □

Lemma 2.6 ([26, Gronwall's lemma with respect to φ function]). *Let $\theta > 0$ and $\varphi \in C^1[J,\mathbb{R}](J=[a,b])$ an increasing function such that $\varphi'(t) \neq 0, \forall t \in J$. Assume that H is a non-negative and non-decreasing, and G is a non-negative function locally integrable on J and suppose also that F is non-negative and locally integrable on J with*

$$F(t) \leqslant G(t) + H(t)\int_a^t \varphi'(\tau)(\varphi(t)-\varphi(\tau))^{\theta_1-1}F(\tau)d\tau,\ t \in J,$$

then, for all $t \in J$,

$$F(t) \leqslant G(t) + \int_a^t \sum_{k=1}^{\infty} \frac{[H(t)\Gamma(\theta)]^k}{\Gamma(k\theta_1)}\varphi'(\tau)(\varphi(t)-\varphi(\tau))^{k\theta_1-1}G(\tau)d\tau.$$

Further, if $G(t)$ be a non-decreasing function on J, then,

$$F(t) \leqslant G(t)E_\theta\left[H(t)\Gamma(\theta)(\varphi(t)-\varphi(a))^{\theta_1}\right],\ t \in J,$$

where $E_\theta(z) = \sum_{j=0}^{\infty}\frac{z^j}{\Gamma(j\theta+1)}$ is called the Mittag-Leffler function.

Theorem 2.7 ([9, Krasnoselskii's fixed point theorem]). *Let χ be a Banach space, let Ω be a bounded closed convex subset of χ and let T_1, T_2 be mapping from Ω into χ such that $T_1x + T_2y \in \Omega$ for every pair $x, y \in \Omega$. If T_1 is contraction and T_2 is completely continuous, then the equation $T_1x + T_2x = x$ has a solution on Ω.*

3. Theoretical results

First of all, we introduce an equivalent integral equation corresponding to the φ-Caputo VF-IDEs given in Eq. (1.1). Let $\chi \in C^{n-1}([c_0, C])$ be a Banach space with the norm $\|\chi\| = \sup_{u\in[c_0,C]}|\chi(u)|$.

Lemma 3.1. *$\chi \in C^{n-1}([c_0, C])$ is a solution of the φ-Caputo VF-IDEs Eq. (1.1) if and only if χ is a solution of the following φ-Riemann-Liouville integral equation*

$$\begin{aligned}\chi(u) = &\sum_{k=0}^{n-1}\frac{\rho_k}{k!}[\varphi(u)-\varphi(c_0)]^k + I^{\theta,\varphi}_{c_0}\Big(\mathfrak{O}\left(u,\chi(u)\right)\Big)\\ &+ I^{\theta,\varphi}_{c_0}\left(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\right) + I^{\theta,\varphi}_{c_0}\left(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\right).\end{aligned} \tag{3.1}$$

Proof. Let $\chi \in C^{n-1}[c_0, C]$ be a solution of Eq. (1.1) . Applying the φ-Riemann-Liouville fractional integral of order θ to both sides of Eq. (1.1), then using Lemma 2.4, we obtain

$$\begin{aligned}\chi(u) &= \sum_{k=0}^{n-1} c_k[\varphi(u)-\varphi(c_0)]^k + \frac{1}{\Gamma(\alpha)}\int_{c_0}^{u}\varphi'(s)(\varphi(u)-\varphi(s))^{\theta-1}\mathfrak{O}\left(s,\chi(s)\right)ds\\ &\quad + I^{\theta,\varphi}_{c_0}\left(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\right) + I^{\theta,\varphi}_{c_0}\left(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\right)\\ &= \sum_{k=0}^{n-1}\frac{\chi^{[k]}_{\varphi}(c_0)}{k!}[\varphi(u)-\varphi(c_0)]^k + I^{\theta,\varphi}_{c_0}\Big(\mathfrak{O}\left(u,\chi(u)\right)\Big)\\ &\quad + I^{\theta,\varphi}_{c_0}\left(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\right) + I^{\theta,\varphi}_{c_0}\left(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\right).\end{aligned} \tag{3.2}$$

Substituting the initial condition $\chi_{\varphi}^{(k)}(c_0) = \rho_k, k = 0, 1, 2, \ldots, n-1$ in the above equation, we obtain Eq. (3.1).

To prove the converse, by applying the fractional φ-Caputo derivative operator ${}^{c}D_{c_0^+}^{\theta,\varphi}$ to both sides of Eq. (3.1) and using Lemma 2.5, we obtain

$$ {}^{c}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = \mathfrak{O}\left(u, \chi(u)\right) + \int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv + \int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv. $$

The initial conditions in Eq. (1.1) are satisfied using direct computations on Eq. (3.2) □

Now, to investigate the existence and the uniqueness of φ-Caputo VF-IDEs given in Eq. (1.1), we construct the following assumptions and operator.

(H_1) There exist the functions $\omega_1, \omega_2, \omega_3 \in C([c_0, C])$ such that $|\mathfrak{H}_1(\chi(u))| \leqslant \omega_1(u)$, $|\mathfrak{H}_2(\chi(u))| \leqslant \omega_2(u)$ and $|\mathfrak{O}\left(u, \chi(u)\right)| \leqslant \omega_3(u)$, for any $\chi \in C^{n-1}([c_0, C])$, $u \in [c_0, C]$. Furthermore, $\omega_1^* = \sup_{u\in J}|\omega_1(u)|$, $\omega_2^* = \sup_{u\in J}|\omega_2(u)|$ and $\omega_3^* = \sup_{u\in J}|\omega_3(u)|$.

(H_2) There exist the constants $\ell_1, \ell_2, \ell_3 > 0$ such that $|\mathfrak{H}_1(\chi_1(u)) - \mathfrak{H}_1(\chi_2(u))| \leqslant \ell_1|\chi_1(u) - \chi_2(u)|$, $|\mathfrak{H}_2(\chi_1(u)) - \mathfrak{H}_2(\chi_2(u))| \leqslant \ell_2|\chi_1(u) - \chi_2(u)|$ and $|\mathfrak{O}\left(u, \chi_1(u)\right) - \mathfrak{O}\left(u, \chi_2(u)\right)| \leqslant \ell_3|\chi_1(u) - \chi_2(u)|$, for any $\chi_1, \chi_2 \in C^{n-1}([c_0, C])$ and $u \in [c_0, C]$.

Define the operator $\Psi : C^{n-1}([c_0, C]) \to C^{n-1}([c_0, C])$ such that

$$ \begin{aligned} \Psi(\chi(u)) = &\sum_{k=0}^{n-1} \frac{\rho_k}{k!}[\varphi(u) - \varphi(c_0)]^k + I_{c_0}^{\theta,\varphi}\left(\mathfrak{O}\left(u, \chi(u)\right)\right) \\ &+ I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\right) + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\right). \end{aligned} \tag{3.3} $$

Theorem 3.2 (Existence). *Suppose that* (H_1) *holds, and*

$$ \sum_{k=0}^{n-1} \frac{\rho_k}{k!}\|\varphi(C) - \varphi(c_0)\|^k + \frac{\|\varphi(C) - \varphi(c_0)\|^\theta}{\Gamma(\theta+1)}(\omega_3^* + \alpha_1\omega_2^* + \alpha_2\omega_2^*) < r, $$

where $r > 0$. *Then, the* φ*-Caputo VF-IDEs, Eq.* (1.1) *has a at least one solution.*

Proof. Let $u \in [c_0, C]$ and $\chi, \tilde{\chi} \in C^{n-1}([c_0, C])$. Using the Krasnoselskii's fixed point theorem, we divide the operator Ψ in Eq. (3.3) into the sum of two operators $\Psi_1 + \Psi_2$ on the compact set $\mathbb{N}_r = \{\chi \in C^{n-1}[c_0, C] : \|\chi\| \leqslant r\}$, where Ψ_1 and Ψ_2 are, respectively, defined by

$$ \Psi_1(\chi(u)) = \sum_{k=0}^{n-1} \frac{\rho_k}{k!}[\varphi(u) - \varphi(c_0)]^k, $$

and

$$ \Psi_2(\chi(u)) = I_{c_0}^{\theta,\varphi}\left(\mathfrak{O}\left(u, \chi(u)\right)\right) + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\right) + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\right). $$

Step1. We prove $(\Psi_1(\chi(u)) + \Psi_2(\tilde{\chi}(u))) \in \mathbb{N}_r$. Using ($H_1$) gives

$$ \begin{aligned} &|\Psi_1(\chi(u)) + \Psi_2(\tilde{\chi}(u))| \\ &\leqslant \sum_{k=0}^{n-1} \frac{\rho_k}{k!}|\varphi(u) - \varphi(c_0)|^k + I_{c_0}^{\theta,\varphi}|\mathfrak{O}\left(u, \tilde{\chi}(u)\right)| + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{C} |\mathfrak{K}_1(u,v)||\mathfrak{H}_1(\tilde{\chi}(v))|dv\right) \end{aligned} $$

$$
\begin{aligned}
&+ I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{u}|\mathfrak{K}_2(u,v)||\mathfrak{H}_2(\tilde{\chi}(v))|dv\right)\\
&\leqslant \sum_{k=0}^{n-1}\frac{\rho_k}{k!}|\varphi(u)-\varphi(c_0)|^k + I_{c_0}^{\theta,\varphi}\left(\omega_3(u)+\int_{c_0}^{C}|\mathfrak{K}_1(u,v)|\omega_1(v)dv+\int_{c_0}^{u}|\mathfrak{K}_2(u,v)|\omega_2(v)dv\right)\\
&\leqslant \sum_{k=0}^{n-1}\frac{\rho_k}{k!}|\varphi(u)-\varphi(c_0)|^k + \frac{(\varphi(u)-\varphi(c_0))^\theta}{\Gamma(\theta+1)}\left(\omega_3(u)+\int_{c_0}^{C}|\mathfrak{K}_1(u,v)|\omega_1(v)dv+\int_{c_0}^{u}|\mathfrak{K}_2(u,v)|\omega_2(v)dv\right).
\end{aligned}
$$

Taking supremum for both sides, we obtain

$$
\begin{aligned}
\|\Psi_1(\chi)+\Psi_2(\tilde{\chi})\| &= \sup_{u\in J}|\Psi_1(\chi(u))+\Psi_2(\tilde{\chi}(u))|\\
&\leqslant \sum_{k=0}^{n-1}\frac{\rho_k}{k!}\|\varphi(C)-\varphi(c_0)\|^k + \frac{\|\varphi(C)-\varphi(c_0)\|^\theta}{\Gamma(\theta+1)}(\omega_3^*+\alpha_1\omega_1^*+\alpha_2\omega_2^*) < r,
\end{aligned}
$$

which implies that $\Psi_1(\chi(u))+\Psi_2(\tilde{\chi}(u))\in\mathbb{N}_r$.

Step2. To prove that Ψ_1 is a contraction on $\mathbb{N}_r$,

$$
\|\Psi_1(\chi)-\Psi_1(\tilde{\chi})\| \leqslant \sum_{k=0}^{n-1}\frac{\rho_k}{k!}|\varphi(u)-\varphi(c_0)|^k - \sum_{k=0}^{n-1}\frac{\rho_k}{k!}|\varphi(u)-\varphi(c_0)|^k = 0.
$$

Hence, Ψ_1 is a contraction with zero Lipschitz constant.

Step3. To prove that Ψ_2 is completely continuous on $\mathbb{N}_r$,

$$
\begin{aligned}
|\Psi_2(\chi(u))| &\leqslant I_{c_0}^{\theta,\varphi}|\mathfrak{O}(u,\chi(u))| + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{C}|\mathfrak{K}_1(u,v)||\mathfrak{H}_1(\chi(v))|dv\right)\\
&\quad + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{u}|\mathfrak{K}_2(u,v)||\mathfrak{H}_2(\chi(v))|dv\right)\\
&\leqslant I_{c_0}^{\theta,\varphi}\left(\omega_3(u)+\int_{c_0}^{C}|\mathfrak{K}_1(u,v)|\omega_1(v)dv+\int_{c_0}^{u}|\mathfrak{K}_2(u,v)|\omega_2(v)dv\right)\\
&\leqslant \frac{(\varphi(u)-\varphi(c_0))^\theta}{\Gamma(\theta+1)}\left(\omega_3(u)+\int_{c_0}^{C}|\mathfrak{K}_1(u,v)|\omega_1(v)dv+\int_{c_0}^{u}|\mathfrak{K}_2(u,v)|\omega_2(v)dv\right).
\end{aligned}
$$

Taking supremum for both sides,

$$
\|\Psi_2(\chi)\| \leqslant \gamma := \frac{\|\varphi(C)-\varphi(c_0)\|^\theta}{\Gamma(\theta+1)}(\omega_3^*+\alpha_1\omega_1^*+\alpha_2\omega_2^*),
$$

which implies that Ψ_2 is bounded, and hence the set $\{\Psi_2(\chi(u)) : \chi\in\mathbb{N}_r\}$ is uniformly bounded. To prove that Ψ_2 is equicontinuous, let $u_1,u_2\in[c_0,C]$, $u_1<u_2$, we have

$$
\begin{aligned}
&|\Psi_2(\chi(u_2))-\Psi_2(\chi(u_1))|\\
&\leqslant I_{c_0}^{\theta,\varphi}\left(|\mathfrak{O}(s,\chi(s))|\right)\Big|_{u_1}^{u_2} + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{C}|\mathfrak{K}_1(u,v)||\mathfrak{H}_1(\chi(v))|dv\right)\Big|_{u_1}^{u_2} + I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{u}|\mathfrak{K}_2(u,v)||\mathfrak{H}_2(\chi(v))|dv\right)\Big|_{u_1}^{u_2}\\
&\leqslant \frac{1}{\Gamma(\theta)}\left(\int_{c_0}^{u_2}(\varphi(u_2)-\varphi(s))^{\theta-1} - \int_{c_0}^{u_1}(\varphi(u_1)-\varphi(s))^{\theta-1}\right)\varphi'(s)|\mathfrak{O}(s,\chi(s))|ds\\
&\quad + \frac{1}{\Gamma(\theta)}\int_{c_0}^{u_2}\varphi'(s)(\varphi(u_2)-\varphi(s))^{\theta-1}\left(\int_{c_0}^{C}|\mathfrak{K}_1(u_2,v)||\mathfrak{H}_1(\chi(v))|dv+\int_{c_0}^{s}|\mathfrak{K}_2(s,v)||\mathfrak{H}_2(\chi(v))|dv\right)ds\\
&\quad - \frac{1}{\Gamma(\theta)}\int_{c_0}^{u_1}\varphi'(s)(\varphi(u_1)-\varphi(s))^{\theta-1}\left(\int_{c_0}^{C}|\mathfrak{K}_1(u_1,v)||\mathfrak{H}_1(\chi(v))|dv+\int_{c_0}^{s}|\mathfrak{K}_2(s,v)||\mathfrak{H}_2(\chi(v))|dv\right)ds+
\end{aligned}
$$

$$\leqslant \frac{[\varphi(u_2)-\varphi(c_0)]^{\theta}}{\Gamma(\theta+1)}\,\omega_3(u_2)-\frac{[\varphi(u_1)-\varphi(c_0)]^{\theta}}{\Gamma(\theta+1)}\,\omega_3(u_1)$$
$$+\frac{1}{\Gamma(\theta)}\int_{c_0}^{u_2}\varphi'(s)(\varphi(u_2)-\varphi(s))^{\theta-1}\left(\int_{c_0}^{C}|\mathfrak{K}_1(u_2,\nu)|\omega_1(\nu)d\nu+\int_{c_0}^{s}|\mathfrak{K}_2(s,\nu)|\omega_2(\nu)d\nu\right)ds$$
$$-\frac{1}{\Gamma(\theta)}\int_{c_0}^{u_1}\varphi'(s)(\varphi(u_1)-\varphi(s))^{\theta-1}\left(\int_{c_0}^{C}|\mathfrak{K}_1(u_1,\nu)|\omega_1(\nu)d\nu+\int_{c_0}^{s}|\mathfrak{K}_2(s,\nu)|\omega_2(\nu)d\nu\right)ds.$$

Taking supremum for both sides, we get

$$\|\Psi_2(\chi(u_2))-\Psi_2(\chi(u_1))\|\leqslant\frac{\|(\varphi(u_2)-\varphi(c_0)\|^{\theta}}{\Gamma(\theta+1)}\left(\|\omega_3^*(u_2)\|+\alpha_1\|\omega_1^*(u_2)\|+\alpha_2\|\omega_2^*(u_2)\|\right)$$
$$-\frac{\|\varphi(u_1)-\varphi(c_0)\|^{\theta}}{\Gamma(\theta+1)}\left(\|\omega_3^*(u_1)\|+\alpha_1\|\omega_1^*(u_1)\|+\alpha_2\|\omega_2^*(u_1)\|\right).$$

Note that, as $u_2\to u_1$, $\|\Psi_2(\chi(u_2))-\Psi_2(\chi(u_1))\|\to 0$. Thus, Ψ_2 is equicontinuous. Hence, $\Psi_2(\mathbb{N}_r)$ is relatively compact on $\in\mathbb{N}_r$. By the Arzelá-Ascoli theorem, Ψ_2 is compact on $\mathbb{N}_r$. Then, by Krasnoselkii theorem, we conclude that Ψ has a fixed point which means that φ-Caputo VF-IDEs, Eq. (1.1) has a at least one solution. □

Theorem 3.3 (Uniqueness). *Let* (H_2) *holds. Then, the φ-Caputo VF-IDEs, Eq.* (1.1) *has a unique solution.*

Proof. Let $\chi,\chi^*\in C^{n-1}([c_0,C])$ be two different solutions of Eq. (1.1), and let $u\in J=[c_0,C]$, using (H_2) we have

$$|\chi(u)-\chi^*(u)|=\left|I_{c_0}^{\theta,\varphi}\left(\mathfrak{O}\left(u,\chi(u)\right)\right)-I_{c_0}^{\theta,\varphi}\left(\mathfrak{O}\left(u,\chi^*(u)\right)\right)+I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{C}\mathfrak{K}_1(u,\nu)\mathfrak{H}_1(\chi(\nu))d\nu\right)\right.$$
$$+I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{u}\mathfrak{K}_2(u,\nu)\mathfrak{H}_2(\chi(\nu))d\nu\right)-I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{C}\mathfrak{K}_1(u,\nu)\mathfrak{H}_1(\chi^*(\nu))d\nu\right)$$
$$\left.-I_{c_0}^{\theta,\varphi}\left(\int_{c_0}^{u}\mathfrak{K}_2(u,\nu)\mathfrak{H}_2(\chi^*(\nu))d\nu\right)\right|$$
$$\leqslant I_{c_0}^{\theta,\varphi}\left[|\mathfrak{O}\left(u,\chi(u)\right)-\mathfrak{O}\left(u,\chi^*(u)\right)|+\left(\int_{c_0}^{C}|\mathfrak{K}_1(u,\nu)||\mathfrak{H}_1(\chi(\nu))-\mathfrak{H}_1(\chi^*(\nu))|d\nu\right)\right.$$
$$\left.+\left(\int_{c_0}^{u}|\mathfrak{K}_2(u,\nu)||\mathfrak{H}_2(\chi(\nu))-\mathfrak{H}_2(\chi^*(\nu))|d\nu\right)\right].$$

Taking the norm we obtain

$$\|\chi-\chi^*\|\leqslant\mu\int_{c_0}^{u}\varphi'(s)(\varphi(u)-\varphi(s))^{\theta-1}\|\chi-\chi^*\|ds,$$

where $\mu:=\frac{\ell_3+\alpha_1\ell_1+\alpha_2\ell_2}{\Gamma(\theta)}$. Applying Gronwall inequality, Lemma 2.6, we have

$$\|\chi-\chi^*\|\leqslant 0+\int_{c_0}^{t}\sum_{k=1}^{\infty}\frac{\aleph\,\Gamma(\theta)]^{k}}{\Gamma\left(k\theta\right)}\psi'(\tau)(\psi(t)-\psi(\tau))^{k\theta-1}\times 0d\tau,$$

which implies $\|\chi-\chi^*\|=0$, and hence $\chi=\chi^*$. Therefore, we deduce that the solution of φ-Caputo VF-IDEs, Eq. (1.1) is unique. □

Next, we investigate the solution's stability of φ-Caputo VF-IDEs, Eq. (1.1), by studying the solution's continuity depends on parameters.

Theorem 3.4. *Suppose that* (H_1) *and* (H_2) *hold, and*

$$\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\Big(\ell_3+\ell_1+\ell_2\Big)<1.$$

Let $\chi(u),\chi^{\epsilon}(u)$ *be the solutions, respectively, of Eq.* (3.1) *and the problem*

$$\begin{cases} {}^{C}\mathfrak{D}_{c_0}^{\theta-\epsilon,\varphi}\chi(u)=\mathfrak{O}\left(u,\chi(u)\right)+\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv+\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv,\\ \chi_{\varphi}^{(k)}(c_0)=\rho_k,\ (u\in J=[c_0,C],c_0\geqslant 0,\ k=0,1,2,\dots,n-1),\end{cases}$$

where $n-1<\theta-\epsilon<\theta<n, n\in\mathbb{N}$. *Then,* $\|\chi^{\epsilon}-\chi\|=\mathcal{O}(\epsilon)$, *for* ϵ *sufficiently small.*

Proof. Let $\chi^{\epsilon}(u)$ be the solution of Eq. (1.1), such that

$$\begin{aligned}\chi^{\epsilon}(u)=&\sum_{k=0}^{n-1}\frac{\rho_k}{k!}[\varphi(u)-\varphi(c_0)]^k+I_{c_0}^{\theta-\epsilon,\varphi}\Big(\mathfrak{O}\left(u,\chi(u)\right)\Big)\\&+I_{c_0}^{\theta-\epsilon,\varphi}\Big(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\Big)+I_{c_0}^{\theta-\epsilon,\varphi}\Big(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\Big).\end{aligned}$$

We have

$$\begin{aligned}|\chi^{\epsilon}(u)-\chi(u)|=&\Big|I_{c_0}^{\theta-\epsilon,\varphi}\Big(\mathfrak{O}\left(u,\chi^{\epsilon}(u)\right)\Big)-I_{c_0}^{\theta,\varphi}\Big(\mathfrak{O}\left(u,\chi(u)\right)\Big)+I_{c_0}^{\theta-\epsilon,\varphi}\Big(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi^{\epsilon}(v))dv\Big)\\&-I_{c_0}^{\theta,\varphi}\Big(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\Big)+I_{c_0}^{\theta-\epsilon,\varphi}\Big(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi^{\epsilon}(v))dv\Big)\\&-I_{c_0}^{\theta,\varphi}\Big(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\Big)\Big|.\end{aligned}\tag{3.4}$$

Adding the terms $\pm I_{c_0}^{\theta-\epsilon,\varphi}\Big(\mathfrak{O}\left(u,\chi(u)\right),\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv,\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\Big)$, then taking the supremum for each part at $u\in J$, we obtain

$$\begin{aligned}I_1&=\|I_{c_0}^{\theta-\epsilon,\varphi}\Big(\mathfrak{O}\left(u,\chi^{\epsilon}(u)\right)\Big)-I_{c_0}^{\theta,\varphi}\Big(\mathfrak{O}\left(u,\chi(u)\right)\Big)\|\\&\leqslant\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\ \ell_3\ \|\chi^{\epsilon}(u)-\chi(u)\|+\|I_{c_0}^{\theta-\epsilon,\varphi}(1)-I_{c_0}^{\theta,\varphi}(1)\|\omega_3^*,\\I_2&=\|I_{c_0}^{\theta-\epsilon,\varphi}\Big(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi^{\epsilon}(v))dv\Big)-I_{c_0}^{\theta,\varphi}\Big(\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\chi(v))dv\Big)\|\\&\leqslant\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\ell_1\|\chi^{\epsilon}(u)-\chi(u)\|+\|I_{c_0}^{\theta-\epsilon,\varphi}(1)-I_{c_0}^{\theta,\varphi}(1)\|\ \omega_1^*,\\I_3&=\|I_{c_0}^{\theta-\epsilon,\varphi}\Big(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi^{\epsilon}(v))dv\Big)-I_{c_0}^{\theta,\varphi}\Big(\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\chi(v))dv\Big)\|\\&\leqslant\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\ \ell_2\ \|\chi^{\epsilon}(u)-\chi(u)\|+\|I_{c_0}^{\theta-\epsilon,\varphi}(1)-I_{c_0}^{\theta,\varphi}(1)\|\ \omega_2^*.\end{aligned}\tag{3.5}$$

Substituting Eqs. (3.5) in the supremum of Eq. (3.4), we obtain

$$\begin{aligned}\|\chi^{\epsilon}(u)-\chi(u)\|\leqslant&\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\ell_3\|\chi^{\epsilon}(u)-\chi(u)\|+\|I_{c_0}^{\theta-\epsilon,\varphi}(1)-I_{c_0}^{\theta,\varphi}(1)\|\ \omega_3^*\\&+\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\ell_1\|\chi^{\epsilon}(u)-\chi(u)\|+\|I_{c_0}^{\theta-\epsilon,\varphi}(1)-I_{c_0}^{\theta,\varphi}(1)\|\ \omega_1^*\end{aligned}$$

$$+\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\ell_2\|\chi^{\epsilon}(u)-\chi(u)\|+\|I_{c_0}^{\theta-\epsilon,\varphi}(1)-I_{c_0}^{\theta,\varphi}(1)\|\ \omega_2^*$$
$$\leqslant \delta_1(\epsilon)\|\chi^{\epsilon}(u)-\chi(u)\|+\delta_2(\epsilon),$$

where

$$\delta_1(\epsilon):=\frac{\|(\varphi(C)-\varphi(c_0)\|^{\theta-\epsilon}}{\Gamma(\theta-\epsilon+1)}\Big(\ell_3+\ell_1+\ell_2\Big)$$

and

$$\delta_2(\epsilon):=\|I_{c_0}^{\theta-\epsilon,\varphi}(1)-I_{c_0}^{\theta,\varphi}(1)\|\ \Big(\omega_3^*+\omega_1^*+\omega_2^*\Big).$$

Hence,

$$\|\chi^{\epsilon}(u)-\chi(u)\|\leqslant\frac{\delta_2(\epsilon)}{1-\delta_1(\epsilon)}=\mathcal{O}(\epsilon).$$

Taking ϵ sufficiently small, $\epsilon\to 0$, then $\|\chi^{\epsilon}-\chi\|=0$, which implies that the solutions $\chi^{\epsilon}=\chi$, means that the stability of the solution for φ-Caputo VF-IDEs, Eq. (1.1), is satisfied, and the proof is completed. □

4. Finite difference-deep learning artificial neural networks (FDL-ANN)

4.1. Description of FDL-ANN

The neural network architecture is similar to the neural system of human brain, consisting of many neural cells under learning algorithm. Mathematically, the main objective of learning neural network that it is similar as minimizing the error function with respect to parameters, then predict the values of these parameters to make it close to the expected training solution. The deep learning neural network architecture consists of four main layers: input-layer, two or more hidden-layers, and output layer.

For general case, let ζ denotes all weights $w_{ji}^{[n]}$, and biases $b^{[n]}$ which are used to construct the neural network, where the superscript $[n]$ refers to the layer's number. The forward propagation starts by feeding forward the input layer with some randomly initial neurons $u=\{u_0,u_1,u_2,\dots,u_k\}$. The Algorithm 1 formulates the process of feed forward the neurons or nod with number k through the input to the output layer using the non-linear, monotonic, and continuously differentiable sigmoid activation function defined by $\Phi=\frac{1}{1+e^{-\lambda u}}$. The result for the output layer can be formulated as $O_k(u)=\Phi_k(w_j^{[n-1]}u+b^{[n-1]})$. Since ζ contains all weights and biases used in this calculation, we can express the final output of the neural network by $\hat{O}(u,\zeta)$ as

$$\hat{O}(u,\zeta)=\sum_{j=1}^{k}w_j^{[n]}O_k(u)+b^{[n]}. \tag{4.1}$$

Next step is to express the approximate neural network train solution of Eq. (1.1) by applying ANN as

$$\hat{\chi}(u,\zeta)=\sum_{\bar{k}=0}^{n-1}a_k u^k+G_n(u)\hat{O}(u,\zeta), \tag{4.2}$$

where $\sum_{\bar{k}=0}^{n-1}a_k u^k$ satisfies the initial conditions, and the second term $G_n(u)\hat{O}(u,\zeta)$ represents the output of the neural network, such that $\hat{O}(u,\zeta)$ is defined in Eq. (4.1), which is used in optimization process in backward-propagation, and $G_n(u)$ is the multiplication function start with degree greater than or equal n which choose to satisfy the stability of the neural network solution.

Substituting Eq. (4.2) into Eq. (1.1) yields

$${}^{\mathcal{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\hat{\chi}(u,\zeta)=\mathfrak{O}\left(u,\hat{\chi}(u,\zeta)\right)+\int_{c_0}^{C}\mathfrak{K}_1(u,v)\mathfrak{H}_1(\hat{\chi}(v,\zeta))dv+\int_{c_0}^{u}\mathfrak{K}_2(u,v)\mathfrak{H}_2(\hat{\chi}(v,\zeta))dv. \tag{4.3}$$

Algorithm 1 Feed forward.

function FEED_FORWARD(u, ζ)
  $w_1, b_1, w_2, b_2, w_3, b_3, \ldots, w_n, b_n \leftarrow$
  $\zeta[w_1], \zeta[b_1], \zeta[w_2], \zeta[b_2], \zeta[w_3], \zeta[b_3], \ldots, \zeta[w_n], \zeta[b_n]$
  $Z_1 \leftarrow$ np.dot$(w_1, u) + b_1$
  $A_1 \leftarrow$ sigmoid(Z_1) ▷ The output of the 1st layer
  $Z_2 \leftarrow$ np.dot$(w_2, A_1) + b_2$
  $A_2 \leftarrow$ sigmoid(Z_2) ▷ The output of the 2nd layer
  $Z_3 \leftarrow$ np.dot$(w_3, A_2) + b_3$
  $A_3 \leftarrow$ sigmoid(Z_3) ▷ The output of the 3rd layer
  Continue until n layers, we have:
  $Z_n \leftarrow$ np.dot$(w_n, A_{n-1}) + b_n$
  $A_n \leftarrow Z_n$ ▷ This is the numerical output
  return A_n ▷ Exactly the value of neural network output is $\hat{O}(u, \zeta) = A_n$
end function

Now, we divide our domain $[c_0, C]$ into k sub-interval, with $u_i = c_0 + ih, h = \frac{C-c_0}{k}, i = 1, 2, \ldots, k$. To reduce the calculation, we use the finite difference method to approximate the φ- Caputo fractional derivative with order $n-1 < \theta < n$, and quadrature approximations to approximate the right-hand side of the equation. For $i = 1, 2, \ldots, k$, we set $\phi_i = \varphi(u_i)$, $\Delta\phi_j = \phi_j - \phi_{j-1}$, and $\hat{\chi}_i = \hat{\chi}(u_i, \zeta)$.

Using L1-type finite difference to the φ-Caputo fractional derivative Definition 2.3, and applying the Taylor expansion on each sub-interval $[u_{j-1}, u_j]$, we obtain

$$\chi_\varphi^{[n]}(s) = \chi_\varphi^{[n]}(u_j) + (s - u_j)\chi_\varphi^{[n+1]}(\xi_s), \quad \xi_s \in [u_{j-1}, u_j]. \tag{4.4}$$

Then, substituting Eq. (4.4) into the integral term of Eq. (2.1), we obtain

$$\int_{u_{j-1}}^{u_j} (\phi_i - \phi(s))^{n-\theta-1} \chi_\varphi^{[n]}(s) d\phi(s) = \chi_\varphi^{[n]}(u_j) \int_{u_{j-1}}^{u_j} (\phi_i - \phi(s))^{n-\theta-1} d\phi(s)(u_j) + R_{i,j}, \tag{4.5}$$

where the remainder is

$$R_{i,j} = \int_{u_{j-1}}^{u_j} (\phi_i - \phi(s))^{n-\theta-1} (s - u_j) \chi_\varphi^{[n+1]}(\xi_s) d\phi(s). \tag{4.6}$$

Hence, the φ-Caputo derivative can be expressed as

$${}^{c}D_{c_0}^{\theta,\varphi}\chi(u_i) \approx \frac{1}{\Gamma(n-\theta)} \sum_{j=0}^{i} A_{i,j}\chi(u_j), \tag{4.7}$$

where the weight-matrix $A_{i,j}$ is given by

$$A_{i,j} = \begin{cases} A_{i,0} = \frac{(\phi_i - \phi_0)^{n-\theta} - (\phi_i - \phi_1)^{n-\theta}}{\Delta\phi_1}, \\ A_{i,j} = \frac{(\phi_i - \phi_{j-1})^{n-\theta} - (\phi_i - \phi_j)^{n-\theta}}{\Delta\phi_j} - \frac{(\phi_i - \phi_j)^{n-\theta} - (\phi_i - \phi_{j+1})^{n-\theta}}{\Delta\phi_{j+1}}, \quad 1 \leqslant j \leqslant i-1, \\ A_{i,i} = (\phi_i - \phi_{i-1})^{n-\theta-1}, \end{cases}$$

with $\phi_j = \varphi(u_j)$. Applying the quadrature approximations for the Volterra-Fredholm integrals and substitute the results with Eq. (4.7) together into Eq. (4.3), we obtain

$$\frac{1}{\Gamma(2-\theta)} \sum_{j=0}^{i} A_{i,j}\hat{\chi}_j = \mathfrak{O}(u_i, \hat{\chi}_i) + h\sum_{j=0}^{k} \mathfrak{K}_1(u_i, u_j)\mathfrak{H}_1(\hat{\chi}_j) + h\sum_{j=0}^{i} \mathfrak{K}_2(u_i, u_j)\mathfrak{H}_2(\hat{\chi}_j). \tag{4.8}$$

This process converts the original equation (4.3) into a system of algebraic equations, which can be solved by Adam optimization techniques that include back-propagation neural networks. The error function at each i is defined by

$$\begin{aligned} E(u_i,\zeta) = \frac{1}{\Gamma(2-\theta)} \sum_{j=0}^{i} A_{i,j}(\hat{\chi}_j;\zeta) - \mathfrak{D}\left(u_i,(\hat{\chi}_i;\zeta)\right) \\ - h\sum_{j=0}^{k} \mathfrak{K}_1\left(u_i,u_j\right)\mathfrak{H}_1\left((\hat{\chi}_j;\zeta)\right) - h\sum_{j=0}^{i} \mathfrak{K}_2\left(u_i,u_j\right)\mathfrak{H}_2\left((\hat{\chi}_j;\zeta)\right). \end{aligned}$$

The Mean Square Error (MSE) is computed as:

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=0}^{N}\left(E(u_i,\zeta)\right)^2.$$

Algorithm 2 represents the pseudocode for optimization process using Adam optimizer.

Algorithm 2 Pseudocode for optimization process using Adam.

Step 1: Initialize the parameters θ, learning rate α, and other Adam hyper-parameters.
Step 2: Define the error function $E(u_i,\zeta)$ for each u_i, then compute the MSE.
Step 3: Compute the gradients of the MSE with respect to the parameters θ.
Step 4: Update parameters using the Adam update rule in Algorithm 3 to reduce the MSE.
for each iteration **do**
  Compute the gradient.
  Update the biased first and second moment estimates.
  Correct the bias.
  Update the parameters.
end for

4.2. *Update the Adam parameters*

We need to update the neural network parameters, ζ using the Algorithm 3.

Algorithm 3 Adaptive moment estimation method.

Input: Initialize parameters $\zeta = [w_1, b_1, w_2, b_2, w_3, b_3, \ldots, w_n, b_n]$, learning rate η, decay rates c_1, c_2, ϵ.
Output: Updated neural network parameters ζ.
Output: Updated neural network parameters ζ.
Initialize: Parameters ζ, first moment $v \leftarrow 0$, second moment $s \leftarrow 0$.
for $l = 1$ to L **do**
  Compute gradient $\frac{\partial E}{\partial \zeta^l}$.
  Update biased first moment estimate: $v^l \leftarrow c_1 v^{l-1} + (1-c_1)\frac{\partial E}{\partial \zeta^l}$.
  Update biased second moment estimate: $s^l \leftarrow c_2 s^{l-1} + (1-c_2)\left(\frac{\partial E}{\partial \zeta^l}\right)^2$.
  Compute bias-corrected first moment estimate: $\hat{v}^l \leftarrow \frac{v^l}{1-c_1^l}$.
  Compute bias-corrected second moment estimate: $\hat{s}^l \leftarrow \frac{s^l}{1-c_2^l}$.
  Update parameters: $\zeta^{l+1} \leftarrow \zeta^l - \frac{\eta}{\sqrt{\hat{s}^l}+\epsilon}\hat{v}^l$.
end for
return ζ

5. Numerical analysis for FDL-ANN method

5.1. Error bound

The main sources of the error when applying finite difference deep learning neural network (FDL-ANN) are the error comes from the neural network approximation error (E_{NN}), finite difference method error (E_{FD}), and the residual minimization error E_R. The neural network approximation error E_{NN} is produced when computed the difference between the approximate solution generated by the deep neural network $\hat{\chi}(u)$ and the actual target solution $\chi(u)$, and by using the universal approximation theorem, for smooth functions on a compact domain, there exists a $\hat{\chi}(u)$ and ϵ_{NN} such that

$$\|\chi - \hat{\chi}\| \leqslant \epsilon_{NN}, \tag{5.1}$$

where $\hat{\chi}(u)$ is defined in Eq. (4.2), which depends on the output network $\hat{O}(u, \zeta)$ defined on (4.1).

For E_{FD}, the truncation error comes when applying the finite difference method on the integration parts in the fractional derivative and the integral terms in the equation. Additionally, the residual minimization error (denoted by E_R) comes from the training process and minimizing the loss function.

Using the results in equations (4.4)-(4.7), for any $s \in [u_{j-1}, u_j]$ with $|s - u_j| \leqslant h$, where h is the step size for finite difference, we have

$$|R_{i,j}| \leqslant \max \left|\chi_\varphi^{[n+1]}\right| \times h \times \int_{u_{j-1}}^{u_j} (\phi_i - \phi(s))^{n-\theta-1} \, d\phi(s) = Ch w_{i,j}, \tag{5.2}$$

where $C = \max_{s\in[c_0,C]} \left|\chi_\varphi^{[n+1]}(s)\right|$. In particular, for the last sub-interval $j = i$,

$$w_{i,i} = \int_{u_{i-1}}^{u_i} (\phi_i - \phi(s))^{n-\theta-1} \, d\phi(s) = \frac{(\phi_i - \phi_{i-1})^{n-\theta}}{n-\theta} \sim (\Delta\phi_i)^{n-\theta},$$

and since $\Delta\phi_i \approx \varphi'(u_i)\, h$, we get

$$|R_{i,i}| = O\left(h\,(\Delta\Phi_i)^{n-\theta}\right) = O\left(h(h)^{n-\theta}\right) = O\left(h^{n+1-\theta}\right),$$

and hence, the local truncation error for the φ-Caputo derivative at $n-1 < \theta \leqslant n$ is

$${}^{C}D_{c_0}^{\theta,\varphi}\chi(u_i) - \frac{1}{\Gamma(n-\theta)} \sum_{j=1}^{i} w_{i,j}\chi_\varphi^{[n]}(u_j) = O\left(h^{n+1-\theta}\right) + \text{ (higher-order)}.$$

The local error on each sub-interval $[u_j, u_{j+1}]$ for the Volterra-Fredholm integral parts in Eq. (4.8) when applying the left-rectangle rule is of order $O\left(h^2\right)$, where the functions inside the integral sign are sufficiently smooth. Summing these errors over $O(1/h)$ intervals gives the total quadrature error of order $O(h)$. Hence, by gathering all errors in the integro-differential equation, we conclude that the error bound of applying the finite difference method with quadrature approximation E_{FD} is of order $O\left(h^{n+1-\theta}\right) + O(h) = O(h)$. The global maximum-norm error satisfies

$$\max_{0\leqslant i\leqslant n} |\chi(u_i) - \chi_i| = \begin{cases} O\left(h^{n+1-\theta}\right), & n = 1, 0 < \theta < 1, \\ O(h), & n \geqslant 2, n-1 < \theta \leqslant n, \end{cases}$$

So, the overall error bound for this method can be summarized as

$$\|\chi - \hat{\chi}\| \leqslant C\left(E_{NN} + E_{FD+R}\right),$$

with C a constant that depends on the stability properties and Lipschitz constants of the operators in the equation.

Theorem 5.1. *Under the assumption* (H_2)*, and letting* $E(u) = \chi(u) - \hat{\chi}(u)$ *be the error function, then, the error involved in the approximation of the* φ*-Caputo type fractional derivative* ${}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0} E(u)$ *is given by*

$$ {}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0} E(u) = \Delta\mathfrak{O}(u) + \Delta I_1(u) + \Delta I_2(u) + E_{FD} $$

with pointwise bound

$$ |{}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0} E(u)| \leqslant \ell \, \|E\| + |E_{FD}|, $$

where ℓ *is non-negative constant and* E_{FD} *is the error producing from finite difference method given in Eq. (5.2). Moreover, the error* $E(u)$ *satisfies*

$$ \|E\| \leqslant C \left(E_{NN} + E_{FD+R} \right), $$

where C *depends on* θ, φ, *and* $\tilde{b}$, *which will be defined later.*

Proof. Now to derive the error bound, we express the pointwise error by $E(u) = \chi(u) - \hat{\chi}(u)$, then we decompose the error into our main equation as follows:

$$ \begin{aligned} {}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0}[\chi(u) - \hat{\chi}(u)] = \mathfrak{O}\left(u, [\chi(u) - \hat{\chi}(u)]\right) + \int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1([\chi(v) - \hat{\chi}(v)])dv \\ + \int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2([\chi(v) - \hat{\chi}(v)])dv + E_{FD}. \end{aligned} $$

After simplifying,

$$ \begin{aligned} {}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0} E(u) = \mathfrak{O}\left(u, [\chi(u) - \hat{\chi}(u)]\right) + \int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1([\chi(v) - \hat{\chi}(v)])dv \\ + \int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2([\chi(v) - \hat{\chi}(v)])dv + E_{FD}. \end{aligned} \tag{5.3} $$

To simplify, we use the notation Δ to indicate to the residual differences. Let $\Delta\mathfrak{O}(u) = \mathfrak{O}\left(u, \chi(u)\right) - \mathfrak{O}\left(u, \hat{\chi}(u)\right)$, $\Delta I_1(u) = \int_{c_0}^{C} \mathfrak{K}_1(u,v)\mathfrak{H}_1([\chi(v) - \hat{\chi}(v)])dv$, $\Delta I_2(u) = \int_{c_0}^{u} \mathfrak{K}_2(u,v)\mathfrak{H}_2([\chi(v) - \hat{\chi}(v)])dv$. Then, Eq. (5.3) can be written as

$$ {}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0} E(u) = \Delta\mathfrak{O}(u) + \Delta I_1(u) + \Delta I_2(u) + E_{FD}. \tag{5.4} $$

Using the Lipschitz continuity mentioned in assumptions (H_2) with $\chi_1(u) = \chi(u)$ and $\chi_2(u) = \hat{\chi}(u)$, and using the conditions in Eq. (1.1) which assume that $\alpha_1 = \max_{u,v\in J} \int_{c_0}^{C} |\mathfrak{K}_1(u,v)|\, dv$, and $\alpha_2 = \max_{u,v\in J} \int_{c_0}^{u} |\mathfrak{K}_2(u,v)|\, dv$, and if we take the supremum, we have

$$ \begin{aligned} \Delta I_1(u) &\leqslant \ell_1 \int_{c_0}^{C} \mathfrak{K}_1(u,v)|E(v)|dv \leqslant \ell_1\alpha_1\|E\|, \\ \Delta I_2(u) &\leqslant \ell_2 \int_{c_0}^{u} \mathfrak{K}_2(u,v)|E(v)|dv \leqslant \ell_2\alpha_2\|E\|, \\ \Delta\mathfrak{O}(u) &\leqslant \ell_3\|E\|. \end{aligned} \tag{5.5} $$

Substituting the results in Eq. (5.5) into Eq. (5.4), we have

$$ |{}^{C}\mathfrak{D}^{\theta,\varphi}_{c_0} E(u)| \leqslant \ell_1\alpha_1\|E\| + \ell_2\alpha_2\|E\| + \ell_3\|E\| + |E_{FD}| \leqslant \ell\|E\| + |E_{FD}|, \tag{5.6} $$

where $\ell = \ell_1\alpha_1 + \ell_2\alpha_2 + \ell_3$, and E_{FD} is the discretization error. Integrating both sides of the above inequality, we note that by using Lemma 2.4, the inverse form for the fractional derivative gives $E(u) =$

$\sum_{k=0}^{n-1} c_k[\varphi(u)-\varphi(c_0)]^k + I_{c_0^+}^{\theta,\varphi}\,{}^{C}\mathfrak{D}_{c_0}^{\theta,\varphi}E(u)$, where $c_k = \frac{\chi_\varphi^{[k]}(c_0)}{k!}$ incorporates the initial errors. Hence, by using Eq. (5.6), we have

$$\begin{aligned}|E(u)| &\leqslant \sum_{k=0}^{n-1} |c_k|\,[\varphi(u)-\varphi(c_0)]^k + I_{c_0^+}^{\theta,\varphi}|{}^{C}\mathfrak{D}_{c_0}^{\theta,\varphi}E(u)| \\ &\leqslant \sum_{k=0}^{n-1} |c_k|\,[\varphi(u)-\varphi(c_0)]^k + \frac{1}{\Gamma(\theta)}\int_{c_0}^{u}(\varphi(u)-\varphi(s))^{\theta-1}\,[\ell\,\|E\|+|E_{FD}|]\,d\varphi(s).\end{aligned}$$

Set $\tilde{a}(u) = \sum_{k=0}^{n-1} |c_k|[\varphi(u)-\varphi(c_0)]^k + \frac{1}{\Gamma(\theta)}\int_{c_0}^{u}(\varphi(u)-\varphi(s))^{\theta-1}\,|E_{FD}(s)|\,d\varphi(s)$, and $\tilde{b} = \frac{\ell}{\Gamma(\theta)}$, then the above inequality can be written as

$$\|E(u)\| \leqslant \tilde{a}(u) + \tilde{b}\int_{c_0}^{u}(\varphi(u)-\varphi(s))^{\theta-1}\|E\|d\varphi(s), \tag{5.7}$$

By applying Grönwall inequality Lemma 2.6 into Eq. (5.7), we obtain

$$\|E(u)\| \leqslant \tilde{a}(u)E_\theta\left(\tilde{b}\,(\varphi(u)-\varphi\,(c_0))^\theta\right), \tag{5.8}$$

where E_θ is the Mittag-Leffler function. Let $\phi(u) = \varphi(u)-\varphi\,(c_0)$, the bounded error in inequality (5.8) becomes

$$\|E(u)\|_{FD+R} \leqslant C_\varphi E_\theta(\tilde{b}\phi(u)) \tag{5.9}$$

where the constant C_φ depends on the transformation function φ. Note from the last inequality that the bounded error is controlled by the total residual error which is a combination of FD error and neural network training error $FD + R$. Additionally, the error expressed in terms of Mittag-Leffler function implies that the error remains bounded under suitable conditions.

Now, combining the total bounded error in Eqs. (5.1) and (5.9), we obtain

$$\|\chi-\hat{\chi}\| \leqslant C_\varphi E_\theta(\tilde{b}\phi(u)) + \epsilon_{NN}.$$

Absorbing constants into a general constant C (which depends on $\tilde{b}, \theta$, and the properties of φ), this is written as

$$\|\chi-\hat{\chi}\| \leqslant C\,(E_{NN}+E_{FD+R})\,,$$

which complete the derivative of the error bound of the proposed method. □

The following two remarks discuss the convergence behavior of FDL-ANN method and the sensitivity of choosing different φ functions.

Remark 5.2. For the convergence behavior of FDL-ANN method, we note that the neural network has a convergence behavior under suitable training conditions such as choosing sufficient network capacity (number of layers), the optimization technique (Adam optimization), and regularization, which makes the approximation error E_{NN} small. While the finite difference method which applied on the integral sign converges depending on the smoothness of the solution. Hence, the combined methods converge such that the residual error (which measures how exactly the neural network satisfies the discretized of the fractional integro-differential equation) is sufficiently reduced during the training process around the order $O\,(\epsilon_{NN}+h^r)$, where r is the FD scheme order.

Remark 5.3. Choosing different φ function plays a crucial role through the definition of φ-Caputo derivative where its including in the kernel form. So, different choices of φ functions such as linear, power-law, exponential or logarithmic functions, will effect on the past history during the training process. For example, the behavior of linear φ function close to classical Caputo derivative, whereas the nonlinear φ can emphasize certain regions over others. Hence, choosing the smooth and well-behaved φ functions give result in more robust and reliable numerical behavior. Additionally, further investigation on the discretization of the finite difference method, we note that if we choose the constant C_φ, which appears in the above error bound, as large as $C_\varphi = \max_{u\in J}\left|\frac{1}{\varphi'(u)}\right|$, it will produce small error that affects on the stability of the proposed method. Moreover, under specific properties of the transformation function φ such as φ' stays bounded away from zero and does not have large oscillations, then the theoretical error bounds remain uniformly controlled, and satisfies the stability of computation.

5.2. Computation complexity

For a uniform discretization finite difference method with N grid points, the time complexity for approximate the fractional derivative, Volterra-Fredholm integrals are generally of order $O(N)$ per evaluation and per grid point when use a direct quadrature rule. Hence, using state-of-the-art techniques such as neural network method with finite difference method can overall the complexity of the computation in the entire system. Moreover, the kernels used in the equation may require $O\left(N^2\right)$ memory, so we approximated them efficiently in the practical implementations to reduce storage needs.

When the neural network method is applied, if we choose the number of parameters m in the network with N collocation points, then time complexity is computed as the cost per epoch of training with $O(N \times m)$. While in practice code, we use CPUs which reduces the effective training time. Also, we choose 1000 epochs in the training process to satisfy the convergence and modified the other hyperparameters. Neural network needs storage of m parameters, and computation of activation function with collocation points in the training process also sorted.

In our hybrid technique, the overall computational effort per epoch is roughly the sum of the costs associated with evaluating the finite difference (FD) approximation and performing a forward pass through the neural network (NN). This gives an overall complexity on the order of $O(N^2 + N \times m)$, where the N^2 term represents the worst-case for the non-local FD computations and $N \times m$ is the cost for the NN evaluation. By employing adaptive grids and fast convolution methods, we reduce the computational cost and time significantly, and overcome the complexity of sorting and ruining the integration parts.

6. Numerical examples

To illustrate the applicability and effectiveness of the presented method, we provided some comparison between φ-Caputo operator with others different operators such as the standard Caputo derivative (Std Caputo), the Atangana-Baleanu derivative (AB), Hilfer derivative (Hilfer). This step might be necessary to understand the numerical performance and stability under each operator, and may help to select the better operator that can fits the real behavior of the modeled system. We provide the following examples.

Example 6.1. Consider the nonlinear fractional integro-differential equation

$$\begin{cases} {}^{\mathfrak{C}}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = \mathfrak{O}\left(u,\chi(u)\right) + \frac{1}{3}\int_0^u (u+v)[\chi(v)]^2 dv + \frac{1}{4}\int_0^1 (u-v)[\chi(v)]^3 dv, & 0 \leqslant u \leqslant 1, \\ \chi_\varphi^{(k)}(0) = 0, & k = 0,1,2, \end{cases} \tag{6.1}$$

where $\theta = 2.2$, $\mathfrak{O}\left(u,\chi(u)\right) = \frac{\Gamma(4)}{\Gamma(1.8)}u^{0.8} - \frac{5}{56}u^8 - \frac{1}{40}u + \frac{1}{44}$. Exact solution is u^3. For $\ell_1 = 2, \ell_2 = \frac{2}{3}$, and for $\ell_3 > 0$ any positive value, and for $\omega_1(u) = u^3$, $\omega_2(u) = u^2$ and $\omega_3(u) = 7 - \frac{5}{56}u^8 - \frac{1}{40}u + \frac{1}{44}$, the conditions in (H_1) and (H_2) are satisfied, and hence this example has existence and unique solution.

The approximate neural network train solution is given by

$$\hat{\chi}(u,\zeta) = u^3\hat{O}(u,\zeta), \tag{6.2}$$

Table 1 demonstrates the comparison with different fractional operators such as the standard Caputo derivative (Std Caputo), the Atangana-Baleanu derivative (AB), Hilfer derivative (Hilfer), and φ-Caputo derivative (φ-Caputo) with different φ functions. Using Adam optimizer with learning rate $\eta = 0.001$, tanh activation function, number of epochs equal 1000, and $\theta = 2.2$. The results show the efficiency of φ-Caputo derivative with different φ functions to give better approximation comparing with Atangana-Baleanu and Hilfer derivatives. The execution time is computed per second for each operator and note that the φ-Caputo operator gives better consuming time during the training process for different φ functions. Figures 1 and 2 show the approximate and error comparison via different fractional operators. The results exhibit the φ-Caputo operator gives good approximation with less error.

Table 1: The comparison for the approximate solutions of various operators comparing with different φ-Caputo, Example 6.1.

u	Exact	Std Caputo	AB	Hilfer	φ-Caputo		
					u^2+u	$\exp(u)$	$\sin(u)$
0.0	0.000	-0.010516	-0.077795	-0.002877	-0.002044	-0.023743	-0.033443
0.1	0.001	-0.010748	-0.083231	-0.002405	-0.001305	-0.025352	-0.036756
0.2	0.008	-0.004936	-0.083035	0.004185	0.005825	-0.020923	-0.034164
0.3	0.027	0.013287	-0.070454	0.023240	0.026039	-0.003867	-0.019408
0.4	0.064	0.050101	-0.038727	0.060867	0.066099	0.032497	0.013468
0.5	0.125	0.111452	0.018658	0.122945	0.132890	0.094780	0.070217
0.6	0.216	0.203290	0.108301	0.215481	0.233598	0.189655	0.156651
0.7	0.343	0.331698	0.237393	0.344679	0.375716	0.324249	0.278673
0.8	0.512	0.502799	0.413451	0.516562	0.566782	0.506043	0.441872
0.9	0.729	0.722526	0.643547	0.736975	0.814346	0.742235	0.651348
1.0	1.000	0.996922	0.935485	1.012149	1.126425	1.040641	0.912523
Time(s)		27.73	15.21	16.77	18.76	14.23	14.38

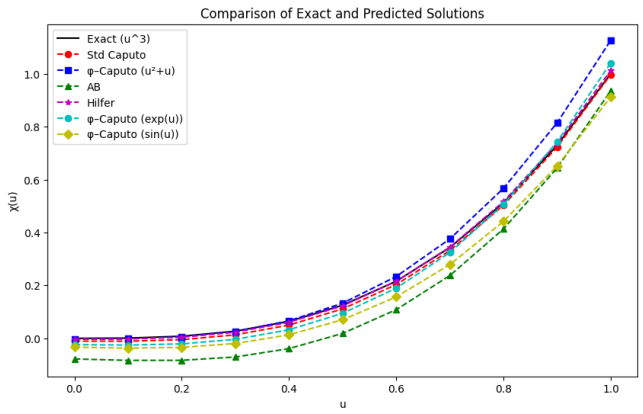

Figure 1: Exact via approximate solution for different operators, Example 6.1.

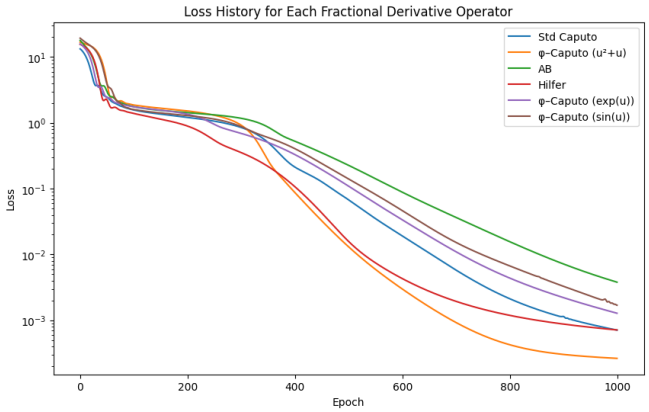

Figure 2: Error function for different fractional operators, 6.1.

Example 6.2. Consider the linear fractional integro-differential equation

$$\begin{cases} {}^{c}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = \mathfrak{O}\left(u,\chi(u)\right) + \frac{1}{4}\int_0^u (u-v)\chi(v)dv + \frac{1}{2}\int_0^1 uv\chi(v)dv, & 0 \leqslant u \leqslant 1, \\ \chi_{\varphi}^{(k)}(0) = 0, & k = 0,1,2, \end{cases} \tag{6.3}$$

where $\theta = 2.3$, and $\mathfrak{O}\left(u,\chi(u)\right) = \frac{\Gamma(4.5)}{\Gamma(2.2)}u^{1.2} - \frac{1}{99}u^{5.5} - \frac{1}{11}u$. Exact solution is $u^{\frac{7}{2}}$. For $\ell_1 = 0.5$, $\ell_2 = \frac{1}{4}$, and for $\ell_3 > 0$ any positive value, and for choosing $\omega_1(u) = 2u$, $\omega_2(u) = 3u$ and $\omega_3(u) = \frac{\Gamma(4.5)}{\Gamma(2.2)}u^{1.2} + \frac{1}{99}u^{5.5} + u$, the conditions in (H_1) and (H_2) are satisfied, and hence this example has existence and unique solution. The approximate neural network train solution is given by

$$\hat{\chi}(u,\zeta) = u^3\hat{O}(u,\zeta). \tag{6.4}$$

Eq. (6.3) is solved by FDL-ANN at different φ-functions, and the results are compared with the Legendre wavelets method (LWM) [22] for $k = 3, M = 2$. Table 2 shows that the absolute errors obtained by FDL-ANN method are lesser than these of LWM. Moreover, the execution time is computed for sigmoid activation function with 1000 epochs. The implementation of this method is easy to use and less costing time for higher order complex equation.

Table 2: The comparison of absolute errors for FDL-ANN at different values of φ-functions with LWM, Example 6.2.

u	FDL-ANN				LWM [22]
	$\varphi = \sin(u)$	$\varphi = \exp(u)$	$\varphi = u$	$\varphi = u^2 + u$	
0.125	4.31607×10^{-4}	4.30394×10^{-4}	4.40295×10^{-4}	4.35207×10^{-4}	2.24750×10^{-4}
0.250	6.85706×10^{-3}	6.82003×10^{-3}	6.91873×10^{-3}	6.89225×10^{-3}	5.68820×10^{-4}
0.375	3.46285×10^{-3}	3.44249×10^{-3}	3.47736×10^{-3}	3.47498×10^{-3}	8.06340×10^{-4}
0.500	1.09332×10^{-3}	1.08667×10^{-3}	1.09514×10^{-3}	1.09608×10^{-3}	$2.37120e\times10^{-3}$
0.625	2.66782×10^{-2}	2.65138×10^{-2}	2.66834×10^{-2}	2.67278×10^{-2}	2.80840×10^{-3}
0.750	5.53006×10^{-2}	5.49596×10^{-2}	5.52603×10^{-2}	5.53785×10^{-2}	3.16120×10^{-3}
0.875	1.02424×10^{-2}	1.01797×10^{-2}	1.02285×10^{-2}	1.02541×10^{-2}	3.35990×10^{-3}
Time(s)	75.17s	60.07s	46.68s	43.05s	

Example 6.3. Consider the nonlinear fractional integro-differential equation

$$\begin{cases} {}^{c}\mathfrak{D}_{c_0}^{\theta,\varphi}\chi(u) = \mathfrak{O}\left(u,\chi(u)\right) + \frac{1}{10}\int_0^u \left(3 + u^2 - v^2\right)\chi(v)dv + \frac{1}{24}\int_0^1 (5+uv)[\chi(v)]^2dv, \\ \chi_{\varphi}^{(0)}(0) = 0, \quad 0 \leqslant u \leqslant 1, \end{cases} \tag{6.5}$$

where $\mathfrak{O}\left(u,\chi(u)\right) = \sqrt{\pi}\left(1+\frac{3u}{4}\right) - \left(\frac{215}{288} + \frac{19u}{180} + \frac{u^{3/2}}{1575}\left(630 + 189u + 120u^2 + 28u^3\right)\right)$, and $\theta = 0.5$. Exact solution is $\sqrt{u}(2+u)$. For $\ell_1 = \frac{1}{6}, \ell_2 = 0.1$ and for $\ell_3 > 0$ any positive value, and for choosing $\omega_1(u) = 2u^2$, $\omega_2(u) = u$ and $\omega_3(u) = \sqrt{\pi}\left(1+\frac{3u}{4}\right) + (1 + u + u^{3/2})\left(630 + 189u + 120u^2 + 28u^3\right)$, then, the conditions in (H_1) and (H_2) are satisfied, and hence this example has existence and unique solution. The approximate neural network train solution is given by

$$\hat{\chi}(u,\zeta) = \hat{O}(u,\zeta). \tag{6.6}$$

Figures 3 and 4 show the results of applying FDL-ANN on Example 6.3, using tanh activation function. Results shows that after epochs 40, loss function will be stable for different fractional derivative, and minimize the error. Moreover, the comparison of these operators gives close values to each other for this example, which indicates to applicable of the proposed method to solve this type of problems.

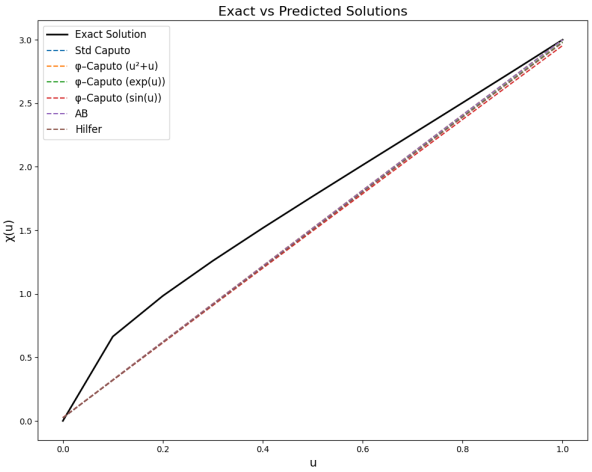

Figure 3: Approximate solutions using different derivatives, Example 6.3.

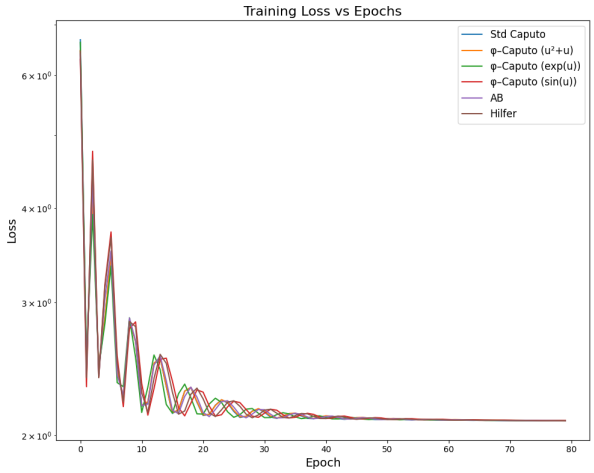

Figure 4: Error functions using different derivatives, Example 6.3.

Example 6.4. Consider the following VFIDEs [14]:

$$\begin{cases} {}_0^C D_t^{\alpha,\varphi}\chi(u) = \mathfrak{O}\,(u,\chi(u)) + \frac{1}{2}\int_0^u e^{-2u}s\sin(\chi(s))ds,\ 1 < \alpha \leqslant 2,\ 0 < t \leqslant 1.5, \\ \chi(0) = 0,\ \chi'_\varphi(0) = 0, \end{cases} \tag{6.7}$$

where the function $\mathfrak{O}\,(u,\chi(u))$ is generated using the exact solution $\chi(u) = (\varphi(u) - \varphi(0))^2$. Figure 5 shows the trail neural network solution with different values of α. The solutions give good approximation when the fractional order close to the integer order 1. Table 3 shows the results of MSE for the extended Chebyshev cardinal wavelets method (EChCWM) [14] and our method FDL-ANN with $\varphi(u) = \exp(u)$. The results of EChCWM is computed at $k = M = 3$, and the neural network solution of degree 3 at $\alpha = 1.25, 1.50$ and 1.75. Clearly, the approximate solutions are comparable by using FDL-ANN as EChCWM, which demonstrates the efficiency of our method for solving the given equation.

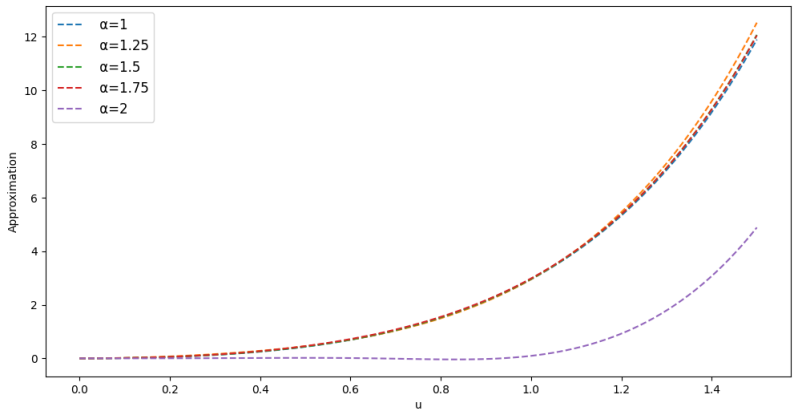

Figure 5: The trail neural network solution at different values of α, Example 6.4.

Table 3: MSE of EChCWM and FDL-ANN methods using $\varphi = \exp(u)$, Example 6.4.

α	EChCWM [14]	FDL-ANN	Time(s) for FDL-ANN
1.25	1.1740×10^{-2}	2.2257×10^{-2}	16.98s
1.50	1.0967×10^{-2}	6.2621×10^{-2}	17.76s
1.75	1.1250×10^{-2}	4.1031×10^{-2}	21.33s

7. Conclusion

In this research, the existence and uniqueness of solution for the higher order φ-Caputo VF-IDEs have been investigated using the Liptchiz conditions, the Krasnoselskii's's fixed point theorem, and the Gronwall inequality with respect to the φ function under some hypotheses. The stability of solution is investigated using the continuity depends on the parameters. Moreover, a new numerical technique which uses the deep learning artificial neural network and finite difference method (FDL-ANN) was developed to solve φ-Caputo VF-IDEs. Feed-forward and Adam optimization methods are used to train the model with one input layer, 3 hidden layers, and one output layer with different activation function. The numerical analysis for the error bound, the complexity of computation are provided. The numerical examples are developed for different values of φ-functions to test the training process, and they demonstrate the efficiency of the proposed method to solve the complicated linear and non-linear higher order FIDEs problems. The FDL-ANN approach provides a good performance for minimizing the errors under the complexity of φ-Caputo operator.

Acknowledgment

The authors would like to thanks the Universiti Putra Malaysia for the financial support through Putra Grant (GP/2023/9752700).

Authors' contribution

Kawthar Alsa'di: writing original draft, visualization, validation, formal analysis, methodology, conceptualization. N. M. A. Nik Long: writing original draft, investigation, validation, supervision, formal analysis, conceptualization, funding acquisition.

References

[1] M. S. Abdo, S. K. Panchal, H. A. Wahash, *Ulam-Hyers-Mittag-Leffler stability for a ψ-Hilfer problem with fractional order and infinite delay*, Results Appl. Math., **7** (2020), 12 pages. 1

[2] S. Abut, H. Okut, K. James Kallail, *Paradigm shift from Artificial Neural Networks (ANNs) to deep Convolutional Neural Networks (DCNNs) in the field of medical image processing*, Expert Syst. Appl., **244** (2024), 16 pages.

[3] M. R. Admon, N. Senu, A. Ahmadian, Z. Abdul Majid, S. Salahshour, *A new efficient algorithm based on feedforward neural network for solving differential equations of fractional order*, Commun. Nonlinear Sci. Numer. Simul., **117** (2023), 27 pages. 1

[4] T. Allahviranloo, A. Jafarian, R. Saneifard, N. Ghalami, S. Measoomy Nia, F. Kiani, U. Fernandez-Gamiz, S. Noeiaghdam, *An application of artificial neural networks for solving fractional higher-order linear integro-differential equations*, Bound. Value Probl., **2023** (2023), 14 pages. 1

[5] R. Almeida, *A Caputo fractional derivative of a function with respect to another function*, Commun. Nonlinear Sci. Numer. Simul., **44** (2017), 460–481. 2.1, 2.2, 2.3, 2.4

[6] R. Almeida, A. B. Malinowska, M. T. T. Monteiro, *Fractional differential equations with a Caputo derivative with respect to a kernel function and their applications*, Math. Methods Appl. Sci., **41** (2018), 336–352. 2.1, 2.2, 2.3, 2.4

[7] M. Amjad, M. Ur Rehman, *A product integration method for numerical solutions of φ fractional differential equations*, J. Comput. Sci., **76** (2024). 1

[8] V. Bohaienko, *Selection of ψ-Caputo derivatives functional parameters in generalized water transport equation by genetic programming technique*, Results Control Optim., **5** (2021), 8 pages. 1

[9] T. A. Burton, C. Kirk, *A fixed point theorem of Krasnoselskii-Schaefer type*, Math. Nachr., **189** (1998), 23–31. 2.7

[10] R. Dechter,*Learning while searching in constraint-satisfaction problems*, University of California, Comput. Sci. Department, Cognitive Systems, (1986), 178–183.

[11] R. Dhayal, Q. Zhu, *Stability and controllability results of ψ-Hilfer fractional integro-differential systems under the influence of impulses*, Chaos Solitons Fractals, **168** (2023), 13 pages. 1

[12] X. Fang, L. Qiao, F. Zhang, F. Sun, *Explore deep network for a class of fractional partial differential equations*, Chaos Solitons Fractals, **172** (2023), 7 pages. 1

[13] V. F. Hatipoğlu, *Forecasting of COVID-19 fatality in the USA: comparison of artificial neural network-based models*, Bull. Malays. Math. Sci. Soc., **46** (2023), 23 pages. 1

[14] M. H. Heydari, M. Razzaghi, *A new wavelet method for fractional integro-differential equations with ψ-Caputo fractional derivative*, Math. Comput. Simul., **217** (2024), 97–108. 1, 6.4, 6.4, 3
[15] A. Jafarian, S. N. Measoomy, *An application of ANNs on power series method for solving fractional Fredholm type integro-differential equations*, Neural Parallel Sci. Comput., **24** (2016), 369–380. 1
[16] U. N. Katugampola, *New fractional integral unifying six existing fractional integrals*, arXiv preprint arXiv:1612.08596 (2016), 6 pages. 2.5, 2
[17] K. Knoerzer, *Leveraging artificial intelligence for simplified adiabatic compression heating prediction: Comparing the use of artificial neural networks with conventional numerical approach*, Innov. Food Sci. Emerg. Technol. **91** (2024), 11 pages.
[18] J. Li, L. Ma, *A unified Maxwell model with time-varying viscosity via ψ-Caputo fractional derivative coined*, Chaos Solitons Fractals, **117** (2023), 12 pages. 1
[19] C. Li, N. G. N'Gbo, F. Su, *Finite difference methods for nonlinear fractional differential equation with ψ-Caputo derivative*, Phys. D, **460** (2024), 10 pages. 1
[20] G. W. Lindsay, *Grounding neuroscience in behavioral changes using artificial neural networks*, Curr. Opin. Neurobiol., **84** (2024).
[21] K. Mukdasai, Z. Sabir, M. A. Z. Raja, R. Sadat, M. R. Ali, P. Singkibud, *A numerical simulation of the fractional order Leptospirosis model using the supervise neural network*, Alex. Eng. J., **61** (2022), 12431–12441. 1
[22] P. Rahimkhani, Y. Ordokhani, E. Babolian, *Fractional-order Bernoulli functions and their applications in solving fractional Fredholem-Volterra integro-differential equations*, Appl. Numer. Math., **122** (2017), 66–81. 6.2, 2
[23] A. Sabir, M. Ur. Rehman, *A numerical method based on quadrature rules for ψ-fractional differential equations*, J. Comput. Appl. Math., **419** (2023), 14 pages. 1
[24] R. Saneifard, A. Jafarian, N. Ghalami, S. M. Nia, *Extended artificial neural networks approach for solving two-dimensional fractional-order Volterra-type integro-differential equations*, Inf. Sci., **612** (2022), 887–897. 1
[25] M. Umar, F. Amin, Q. Al-Mdallal, M. R. Ali, *A stochastic computing procedure to solve the dynamics of prevention in HIV system*, Biomed. Signal Process. Control, **78** (2022).
[26] J. Vanterler da Costa Sousa, E. Capelas de Oliveira, *A Gronwall inequality and the Cauchy-type problem by means of ψ-Hilfer operator*, Differ. Equ. Appl., **11** (2019), 87–106. 2.6
[27] H. Vu, N. Van Hoa, *Hyers-Ulam stability of fuzzy fractional Volterra integral equations with the kernel ψ-function via successive approximation method*, Fuzzy Sets Syst., **419** (2021), 67–98.
[28] J.-L. Zhou, S.-Q. Zhang, Y.-B. He, *Existence and stability of solution for nonlinear differential equations with ψ-Hilfer fractional derivative*, Appl. Math. Lett., **121** (2021), 7 pages. 1
[29] C. J. Zúñiga-Aguilar, H. M. Romero-Ugalde, J. F. Gómez-Aguilar, R. F. Escobar-Jiménez, M. Valtierra-Rodríguez, *Solving fractional differential equations of variable-order involving operators with Mittag-Leffler kernel using artificial neural networks*, Chaos Solitons Fractals, **103** (2017), 382–403. 1