

## 整数階でない微積分とニューラルネットワーク

村瀬 一之<sup>† a)</sup>

Calculus of Non-Integer Order and Neural Network (Japanese Edition)

Kazuyuki MURASE<sup>† a)</sup>

**あらまし** 一般に広く用いられている微分や積分では階数は整数とされている。例えば、1階微分  $dy/dx$ 、2階微分  $d^2y/dx^2$ 、1階積分  $\int ydx$ 、である。この階数が実数値をとる非整数階の微積分法は数百年も前から数学的には提案されている。非整数階の微積分法は拡散や履歴現象を記述するのに適していることから、粘弾性物質や流体の動態などを表す際には用いられてきた。昨今、人工知能に用いられるニューラルネットワークの学習過程などに非整数階の微分を用いるとその性能が上がることが知られ、さらに最近、ディープラーニングにも用いた例が報告されている。ここではそれらの事例を簡単に紹介する。

**キーワード** 非整数階微積分, ニューラルネットワーク, 最急降下法, 誤差逆伝播学習, 畳み込みニューラルネットワーク, 深層学習

**Abstract** In widely used differentiation and integration, the order is assumed to be an integer, for example, the first-order derivative  $dy/dx$ , second-order derivative  $d^2y/dx^2$ , first-order integral  $\int ydx$ , etc. Calculus of non-integer order has been built mathematically before hundreds of years. It is suitable for describing dispersion and hysteresis. Therefore, it is traditionally used to express dynamics of viscoelastic materials and fluids. Recently, it is highlighted that differentials of non-integer order can be effectively used in the learning process of neural networks which are the primary component of deep learning in the latest artificial intelligence. This brief article describes some latest progress.

**Keywords** Fractional calculus, Neural network, Gradient descent, Back-propagation learning, Convolutional neural network, Deep learning

### 1. まえがき

非整数階の微積分法は整数でない階数の微分や積分を伴う数学である。1695年に G.W. Leibniz が初めて記述し、1823年の N.H. Abel の文献にほとんどの要素がみられるとされている [1]。まとまった文献は 1974年の Oldham & Spanier のものとされる [2]。

整数階の微分が局所的な性質を表すのに対し、非整数階の微分は履歴を含むことから、様々な拡

散や履歴現象の記述に適しているとされる [2][3]。すなわち、粘弾性を持つ物質、電磁波の拡散、熱伝導といった自然現象である。

近年、非整数階の微分を使ったニューラルネットワーク (NN) の最急降下法や誤差逆伝播学習が脚光を浴びている。300編を超える参考文献を引用した総説にその詳細が記されている [4]。再帰型 NN では、非線形システム同定、パタン認識、Mackey-Glass カオス時系列予測などが行われている。順伝播型 NN はクラス分類、時系列分析や各種の制御に使われている。いずれにおいても、整数階の微分を使ったモデルより性能が高いことが示されている。

本小文では、まず非整数階の微積分について簡単に説明し、NN の誤差逆伝播学習に使われた例

<sup>†</sup> 大阪国際工科専門職大学 工科学部

Faculty of Information Technology, International Professional University of Technology in Osaka, 3-3-1 Umeda, Kita-ku, Osaka 530-0001, Japan

a) E-mail: murase@u-fukui.ac.jp

を示し、さらに畳み込み NN に使われた例を示す。また、その他の応用についても触れる。これらを通じ、非整数階の微積分特性を持つ NN の理解が促進され、各種の応用においてその高性能化の一助とすることを旨とする。

## 2. 非整数階の微積分法

非整数階の微積分法は、英語では Fractional Calculus とか Calculus of non-integer order と呼ばれ、日本語でも分数階微積分学とも呼ばれる。詳しい解説は他に良書があるのでそれらを参考に検索されたい [2][3]。ここでは先に進めるために数学的厳密さは求めずに形式的に説明する。

まずは積分演算オペレータ  $I$  を考える。以下、積分の下限は 0 でなくてもよいが注意を要する。

$$If(t) = \int_0^t f(\tau) d\tau$$

これを 2 回繰り返したものは 1 階積分に帰着できる。

$$I^2 f(t) = \int_0^t d\tau_1 \int_0^{\tau_1} f(\tau) d\tau = \int_0^t (t-\tau) f(\tau) d\tau$$

さらに  $n$  階積分も、Cauchy の反復積分についての公式によって 1 階積分に帰着する。

$$\begin{aligned} f(t) &= \int_0^t d\tau_{n-1} \int_0^{\tau_{n-1}} d\tau_{n-2} \cdots \int_0^{\tau_1} f(\tau) d\tau \\ &= \frac{1}{(n-1)!} \int_0^t (t-\tau)^{n-1} f(\tau) d\tau \end{aligned}$$

ガンマ関数  $\Gamma(\lambda) = \int_0^\infty t^{\lambda-1} e^{-t} dt$  は  $\lambda$  が正整数の時  $(\lambda-1)!$  になるので、ガンマ関数を用いてこの式を非整数  $\lambda$  に拡張する。これが非整数階積分である。

$$I^\lambda f(t) = \frac{1}{\Gamma(\lambda)} \int_0^t (t-\tau)^{\lambda-1} f(\tau) d\tau$$

適当な条件下では、

$$I^{\lambda_1} I^{\lambda_2} = I^{\lambda_2} I^{\lambda_1} = I^{\lambda_1 + \lambda_2}, \lambda_1 > 0, \lambda_2 > 0$$

積分の階数が負の場合については、非整数階微分と呼ばれる。すなわち、 $\lambda > 0$  のとき  $\lambda$  階の非整数微分は、

$$I^{-\lambda} f(t) \equiv D_t^\lambda f(t) \equiv \frac{d^\lambda f(t)}{dt^\lambda} = \frac{1}{\Gamma(-\lambda)} \int_0^t (t-\tau)^{-\lambda-1} f(\tau) d\tau$$

適当な条件下では、

$$D^{\lambda_1} D^{\lambda_2} = D^{\lambda_2} D^{\lambda_1} = D^{\lambda_1 + \lambda_2}, \lambda_1 > 0, \lambda_2 > 0$$

ここで、 $m < \lambda < m + 1$  を満たす整数  $m$  を導入する

と  $0 < \lambda - m < 1$  となり、

$$\begin{aligned} D_t^\lambda f(t) &= D_t^{\lambda-m} D_t^m f(t) \\ &= \frac{1}{\Gamma(m-\lambda)} \int_0^t (t-\tau)^{m-\lambda-1} f^{(m)}(\tau) d\tau \end{aligned}$$

ただし、

$$D_t^m f(t) \equiv \frac{d^m f(t)}{dt^m} \equiv f^{(m)}(t)$$

これは Caputo の非整数階微分 (Caputo fractional-order derivative) と呼ばれる。

他にも非整数階微分の表現があり、適当な条件下でそれぞれ使い分けられている。

Riemann-Liouville (RL) fractional-order derivative では  $m$  階微分が積分の外にある。

$$\begin{aligned} {}^{RL}D_t^\lambda f(t) &= \frac{d^m}{dt^m} D_t^{\lambda-m} f(t) \\ &= \frac{1}{\Gamma(m-\lambda)} \frac{d^m}{dt^m} \int_0^t (t-\tau)^{m-\lambda-1} f(\tau) d\tau \end{aligned}$$

Grunward-Letnikov (GL) fractional-order derivative は積分の下限が明示的である。

$$\begin{aligned} {}^{GL}D_t^\lambda f(t) &= \frac{1}{\Gamma(m-\lambda+1)} \int_a^t (t-\tau)^{m-\lambda} f^{(m+1)}(\tau) d\tau \\ &\quad + \sum_{k=0}^m \frac{(t-a)^{-\lambda+k}}{\Gamma(-\lambda+k+1)} f^{(k)}(a) \end{aligned}$$

簡単な例は次のとおりである。

$$f(x) = 1 \text{ では, } D_x^{1/2} f(x) = 2\pi^{-1/2} x^{-1/2} = 1/\sqrt{\pi x}$$

$$f(x) = x \text{ では, } D_x^{1/2} f(x) = 2\pi^{-1/2} x^{1/2} = 2\sqrt{x/\pi}$$

$$f(x) = \sqrt{x} \text{ では, } D_x^{1/2} f(x) = \sqrt{\pi}/2$$

$$f(x) = e^{ax} \text{ では, } D_x^\lambda f(x) = a^\lambda e^{ax}$$

ガウス関数の非整数階の微分について、

$f(x) = e^{-x^2}$  の場合、 $D_x^\nu f(x)$  をいろいろな非整数階  $\nu$  について示したものが図 1 である。

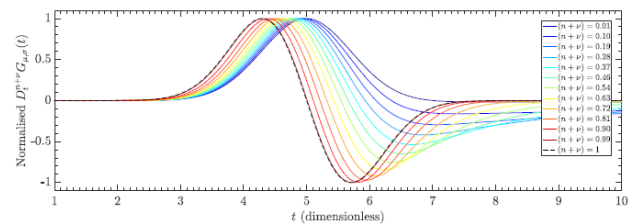


図 1  $D_x^\nu e^{-x^2}$  をいろいろな  $\nu$  について求めたもの ( $n=0$ )  
非整数階微分では長く尾を引く。G. Jorge M. Cruz-Duarte, Juan Rosales-Garcia, C. Rodrigo Correa-Cely, Arturo Garcia-Perez, Juan Gabriel Avina-Cervantes, "A closed form expression for the Gaussian-based Caputo-Fabrizio fractional derivative for signal processing applications," Communications in Nonlinear Science and Numerical Simulation, 61, 138-148, 2018 [5] の Fig.1(a) を転載。

Figure 1 Behavior of  $D_x^\nu e^{-x^2}$  for various non-integer values  $\nu$  ( $n=0$ )  
Excerpted from Fig.1(a) of ref. [5].

### 3. 非整数階微分を用いた誤差逆伝播学習

ニューラルネットワーク (NN) は生物の神経系の仕組みをまねた計算モデルで, 人工知能 (AI) に広く用いられている. その一番の特徴は, 与えられたデータによってその性質を徐々に変化させ, 何らかの評価関数に基づいた最適値に近づけることができる点である. すなわち, 何かを獲得 (学習) させることができる.

まず, Wang et al. (2017) による最急降下法による誤差逆伝播学習 [6] を紹介する. 入力層, 中間層, 出力層の 3 層にそれぞれ  $p, q, 1$  個のニューロン (素子) があり, 入力層から中間層, 中間層から出力層に向けて結合がある 3 層順伝搬型 NN を考える. 入力層から中間層への結合荷重を  $\mathbf{v}_i = (v_{i1}, v_{i2}, \dots, v_{ip})^T$ ,  $i = 1, 2, \dots, q$ , 中間層から出力層への荷重を  $\mathbf{u} = (u_1, u_2, \dots, u_q)^T$  とし, 荷重マトリクスを次のように定義する.

$$\mathbf{w} = (\mathbf{u}, \mathbf{V}) = (\mathbf{u}^T, \mathbf{v}_1^T, \mathbf{v}_2^T, \dots, \mathbf{v}_q^T)^T$$

中間層の活性化関数を  $G(z) = (g(z_1), g(z_2), \dots, g(z_n))^T$ , 出力層の活性化関数を  $y = f(x)$  とすると,  $j$  番目の入力  $x^j$  が与えられたときの出力は  $y = f(\mathbf{u} \cdot G(\mathbf{V}\mathbf{x}^j))$  で表される.

評価関数を二乗誤差とすると, 望ましい出力 (教師信号)  $O^j$  に対する誤差は,

$$E(\mathbf{w}) = \frac{1}{2} \sum_{j=1}^J (O^j - y)^2$$

原著では Rumelhart et al. (1986) による古典的な誤差逆伝播学習 [7] をもとに, 非整数階微分を用いた誤差逆伝播学習について詳細な展開がなされている. 以下には, その概念を要約したものを示す.

古典的な誤差逆伝播学習では, 結合荷重は,  $k$  回目から  $k + 1$  回目の繰り返しの時に, 一階微分を用いた次のような式で更新される.

$$w_{\xi}^{k+1} = w_{\xi}^k - \eta \frac{\partial E(\mathbf{w})}{\partial w_{\xi}}$$

ここで,  $\eta$  は学習率,  $\xi$  は各層間の荷重の一つを表す.

原著ではこの更新に非整数階である  $\alpha$  階微分を用いる. ただし,  $0 < \alpha < 1$  であるが, 前節の議論から  $(m + \alpha)$  階に一般化できると考えられる.

$$w_{\xi}^{k+1} = w_{\xi}^k - \eta D_{w_{\xi}}^{\alpha} E(\mathbf{w})$$

活性化関数  $g, f$  がシグモイド関数, 荷重が有界で有限個, 学習率は正で有界, などの条件を満たすとき, このネットワークが収束することを証明している.

まず, 一般的なデータセットを学習させて, 3 つの非整数階微分, Caputo, GL, RL, を比較している. Iris, Liver, Sonar の 3 つのデータセットを用いた数値実験で, トレーニングデータとテ

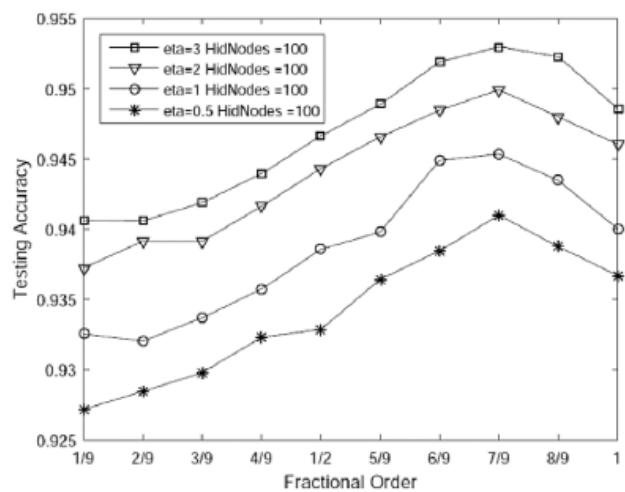
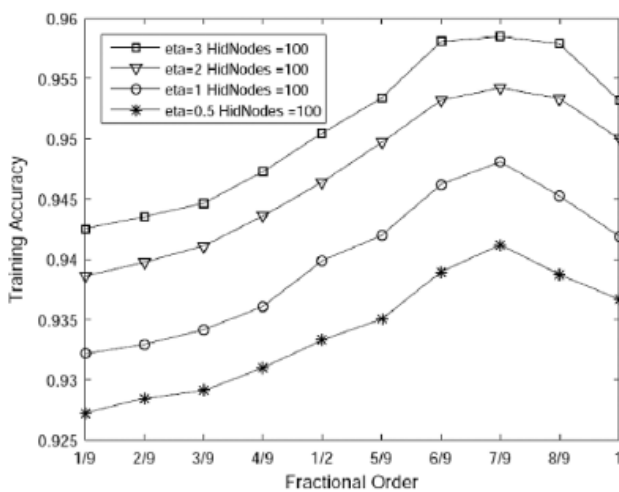


図2 トレーニングとテスト精度を種々の学習率について, 非整数階の微分の階数に対してプロットしたもの

Wang, Y. Wen, Y. Gou, Z. Ye, H. Chen, "Fractional-order gradient descent learning of BP neural networks with Caputo derivative," *Neural Networks*, Vol.89, pp.19-30, 2017 [6] の Fig.2 より一部を転載.

Figure 2 Training and test accuracies for various learning rates are plotted against non-integer fractional order. Excerpted from Fig.2 of ref. [6].

ストデータでの精度は、Caputo 微分が比較的高いが、計算時間は RL 微分が他の約半分となっている。

次に、少し規模の大きいデータセットの学習として、手書き文字認識のデータセット MNIST を用いて、中間層の素子数などをいろいろと変化させて実験している。その結果の一例を図2に示す。非整数階 7/9 付近でトレーニングとテストの両方の精度が最大となっている。他の条件でも同様の傾向がみられる。収束も古典的な誤差逆伝播学習よりも早く安定している。

MNIST データセットを Caputo の非整数階微分を使った誤差逆伝播を進化手法を組み合わせる学習させる試みを Chen et al. (2020) も報告しているが [8]、やはり整数階よりも良い結果を得ている。Pu & Wang (2019) は GL の非整数階微分を使ってより一般的な誤差逆伝播学習について、特にその収束性を検討している [9]。Wang et al. (2022)

は RL の非整数階微分を用いて、より早い精度の高い最急降下法を提案し、誤差逆伝播学習や畳み込みニューラルネットワークの例を示している [10]。Pu et al. (2015) は GL の非整数階微分を用いて、最急降下法を提案しているが [11]、非整数階微分を用いた場合の最適点は整数階を用いた場合の最適点とは異なることを指摘している。

#### 4. 非整数階微分を使った畳み込みニューラルネットワーク

Shen et al. (2020) は畳み込みニューラルネットワーク (Convolutional Neural Network; CNN) に Caputo の非整数階微分を用いた最急降下法による誤差逆伝播学習を適用している [12]。微分の連鎖律 (チェーンルール) を簡単化するために、層間の逆伝播に 1 階の微分を導入している。図3にその逆伝播のスキームを示す。実際に図4に示す多層の CNN である LeNet-5 で手書き文字データ MNIST を学習させている。その結果、収束に必要な時間は整数階のものとはほぼ同じで、階数が 0.9 や 1.1 でトレーニングとテストの精度が 1 階のものより高いことを示している。

Zhou et al. (2020) はロボットマニピュレータのターミナルスライディングモード (TSM) による追従制御を Caputo の非整数階微分と 2 層の畳み込み層を持つ CNN で行い、従前のものに比べて外乱やノイズに対して頑強であることを示している [13]。

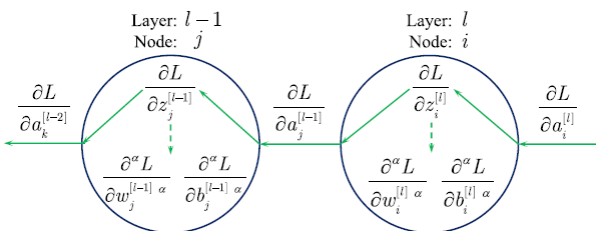


図3 非整数階微分と 1 階微分を組み合わせた逆伝播

D. Sheng, Y. Wei, Y. Chen, and Y. Wong, "Convolutional neural networks with fractional order gradient method," Neurocomputing, Vol.408, pp.42-50, 2020 [12] の Fig.3 より転載。

Figure 3 Backpropagation where non-integer order differentials are combined with first-order differentials. Excerpted from Fig. 3 of ref. [12].

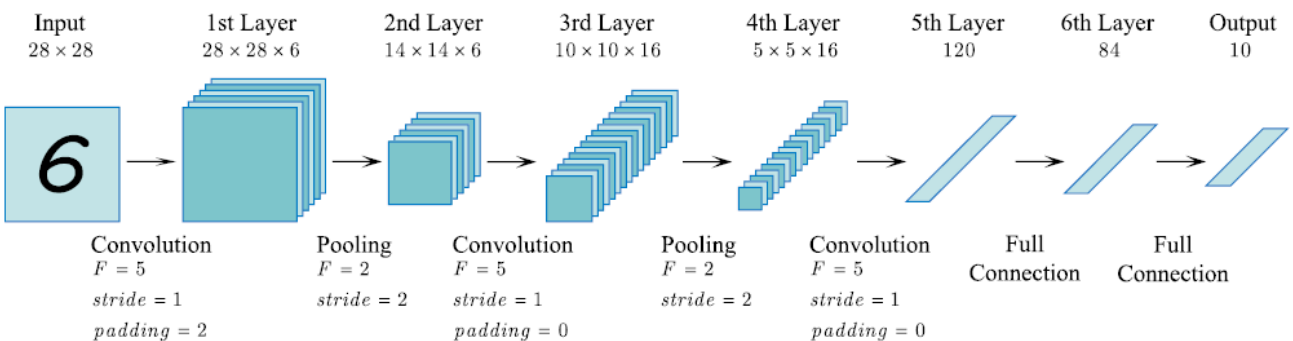


図4 畳み込みニューラルネットワーク LeNet-5.

D. Sheng, Y. Wei, Y. Chen, and Y. Wong, "Convolutional neural networks with fractional order gradient method," Neurocomputing, Vol.408, pp.42-50, 2020 [12] の Fig.5 より転載。

Figure 4 Convolutional neural network LeNet-5. Excerpted from Fig.5 of ref. [12].



## 5. その他の応用

近年、非整数階微分を使ったニューラルネットワークは、医療、暗号化、画像処理、ロボティクス、などに幅広く応用されている。古典的な応用としては、カオス、関数近似、熱伝導過程、周期性、散逸性、などの解析がある。これらの文献は Viera-Martin et al. (2022) による総説 [4] に詳しい。最近の応用としては、Khan et al. (2022) による推薦システム [14], Yu et al. (2022) による画像の暗号化 [15], Singh et al. (2021) による株価の予測 [16], などがある。また、Humphries et al. (2020) [17], Zhang et al. (2021) [18], Li et al. (2022) [19] など、近年、四元数ニューラルネットワークに非整数階微分を導入した研究が多数ある。

## 6. むすび

本小文では非整数階微分のニューラルネットワークへの適用を概観した。非整数階微分はその定義が積分で表されていることからわかるように、局所の性質のみではなく、履歴などといった情報が含まれる。このような性質をもつ現象の分析や表現には最適と考えられる。種々の文献によると計算的な負荷はさほど大きくはないことが示されている。将来的な応用としては、整数階の微分を用いているものを非整数階に置き換えて精度や速度を検証するのみならず、脳波や筋電、認知過程などといった多次元で非線形な現象の解析に四元数とともに用いると効果的と考えられる。

## 文 献

- [1] Fractional Calculus (13:30 JST, December 14, 2022), the free encyclopedia Wikipedia English version, [https://en.wikipedia.org/wiki/Fractional\\_calculus](https://en.wikipedia.org/wiki/Fractional_calculus)
- [2] 杉本信正 “整数階でない微積分法について” 日本数学会 2016 年度秋季総合分科会市民講演会資料, 2016 年 9 月 19 日
- [3] 杉本信正 “非整数階微分・積分とその応用” 日本流体力学会誌「ながれ」, Vol.4, pp.110-120, 1985.
- [4] E. Viera-Martin, J. F. Gomez-Aguilar, J. E. Solis-Perez, J. A. Hernandez-Perez, and R. F. Escobar-Jimenez, “Artificial neural networks: a practical review of applications involving fractional calculus,” *European Physical Journal, Special Topics*, Vol.231, pp.2059–2095, 2022.
- [5] M. Cruz-Duarte, Juan Rosales-Garcia, C. Rodrigo Correa-Cely, Arturo Garcia-Perez, Juan Gabriel Avina-Cervantes, “A closed form expression for the Gaussian-based Caputo-Fabrizio fractional derivative for signal processing applications,” *Communications in Nonlinear Science and Numerical Simulation*, 61, 138-148, 2018.
- [6] J. Wang, Y. Wen, Y. Gou, Z. Ye, H. Chen, “Fractional-order gradient descent learning of BP neural networks with Caputo derivative,” *Neural Networks*, Vol.89, pp.19-30, 2017.
- [7] D.E. Rumelhart, G.E. Hinton, and R.J. Williams, “Learning representations by back-propagating errors,” *Nature*, Vol.323, pp.533–536, 1986.
- [8] M.R. Chen, B.P. Chen, G.Q. Zeng, K.D. Lu, and P. Chu, “An adaptive fractional-order BP neural network based on extremal optimization for handwritten digits recognition,” *Neurocomputing*, Vol.391, pp.260-272, 2020.
- [9] Y.F. Pu and J. Wang, “Fractional-order backpropagation neural networks trained by improved fractional-order steepest descent method, arXiv: 1906.09524, 2019.
- [10] Y. Wang, Y. He, and Z. Zhu, “Study on fast speed fractional order gradient descent method and its application in neural networks,” *Neurocomputing*, Vol.489, pp.366-376, 2022.
- [11] Y.F. Pu, J.L. Zhou, Y. Zhang, N. Zhang, G. Huang, and P. Siarry, “Fractional extreme value adaptive training method: Fractional steepest descent approach,” *IEEE Transactions on Neural Networks and Learning Systems*, Vol.26, No.4, April 2015.
- [12] D. Sheng, Y. Wei, Y. Chen, and Y. Wong, “Convolutional neural networks with fractional order gradient method,” *Neurocomputing*, Vol.408, pp.42-50, 2020.
- [13] M. Zhou, Y. Feng, C. Xue, and F. Han, “Deep convolutional neural network based fractional-order terminal sliding-mode control for robotic manipulators,” *Neurocomputing*, Vol.416, pp.143-151, 2020.
- [14] Z.A. Khan, N.I. Chaudhary, M.A.Z. Raja, “Generalized fractional strategy for recommender systems with chaotic ratings behavior,” *Chaos, Solutions and Fractals*, Vol.160, 112204, 2022.
- [15] F. Yu, X. Kong, H. Chen, Q. Yu, S. Chai, Y. Huang, and S. Du, “A 6D fractional-order memristive Hopfield neural network and its application in image encryption,” *Frontiers in Physics*, 10:847385, 2022.
- [16] N. Singh, Sugandha, T. Mathur, S. Agarwal, and K. Tiwari, “Stock price prediction using fractional gradient-based long short term memory,” *Journal of Physics: Conference Series, IRMAS2021*, 1969:012038, 2021.
- [17] U. Humphries, G. Rajchakit, P. Kaewmesri, P. Chanthorn, R. Sriraman, R. Samidurai, C.P. Lim, “Global stability analysis of fractional-order quaternion-valued bidirectional associative memory neural networks,” *Mathematics*, Volume 8, Issue 5, Article number 801, 1 May 2020.
- [18] W. Zhang, C. Sha, J. Cao, G. Wang, and Y. Wang, “Adaptive quaternion projective synchronization of fractional order delayed neural networks in quaternion field,” *Applied Mathematics and Computation*, Vol.400, 126045, 2021.
- [19] R. Li, and J. Cao, “Passivity and dissipativity of fractional-order quaternion-valued Fuzzy memristive neural networks: Nonlinear

scalarization approach,” IEEE Transactions on Cybernetics,  
Vol.52, No.5, pp.2821-2832, May 2022.



村瀬 一之

米国アイオワ州立大学・院博士。知能情報学の研究に従事



この記事は Creative Commons 4.0 に基づきライセンスされます  
(<https://creativecommons.org/licenses/by/4.0/deed.ja>)。