Journal of Interdisciplinary Mathematics
ISSN: 0972-0502 (Print), ISSN: 2169-012X (Online)
Vol. 29 (2026), No. 3, pp. 701–708
DOI : 10.47974/JIM-2505

Fixed-point theorems in training deep neural networks

Praniti Ram Patil [†]
Department of Information Technology
Dwarkadas Jivanlal Sanghvi College of Engineering
Mumbai 400056
Maharashtra
India

Suresh Singh Rajpurohit *
Department of Artificial Intelligence and Machine Learning
Thakur Shyamnarayan Engineering College
Kandivali
Mumbai 400101
Maharashtra
India

Manthan Joshi [§]
Department of Information Technology
Shree L.R. Tiwari College of Engineering
Thane 401107
Maharashtra
India

Harsha Dave [‡]
Department of Computer Engineering
Shree L. R. Tiwari College of Engineering
Thane 401107
Maharashtra
India

[†] *E-mail:* pranitirpatil@gmail.com
* *E-mail:* srajpurohit422@gmail.com (Corresponding Author)
[§] *E-mail:* manthan.joshi@slrtce.in
[‡] *E-mail:* daveharsha.dave@gmail.com

© Taru Publications

Aakash Mansaram Gojare [@]
Department of Computer Engineering
Shree L. R. Tiwari College of Engineering
Thane 401107
Maharashtra
India

Ashwini Prasad Parkar [#]
Department of Computer Science and Engineering (Data Science)
Parshvanath Charitable Trust's A. P. Shah Institute of Technology
Thane 400615
Maharashtra
India

Abstract

This paper discuss fixed-point theories can be used to train deep neural networks. We find a mathematical link between nonlinear analysis and learning dynamics by looking at neural network optimisation as a fixed-point problem. In repeated training processes, fixed-point formulas ensure convergence, stability, and equilibrium in a theoretical sense. The results of experiments show that fixed-point–based algorithms are faster and more stable than standard optimisation methods such as Adam and gradient descent. This study shows how fixed-point principles can help make learning methods for complicated neural systems more accurate, easy to understand, and effective.

Subject Classification: *37C25, 68T07.*

Keywords: *Fixed-Point theorems, Deep learning, Convergence, Stability, Optimization.*

I. Introduction

Fixed points are very important in mathematical analysis because they help us understand how stable states work in processes that repeat. In deep learning, training neural networks can be seen as a process that tries to find this kind of balance, where the model's parameters stay the same even after many optimisation changes [1]. By connecting gradient-based updates to fixed-point rounds, one can learn more about how training works, how stable optimisation is, and how generalisation works [2, 3]. This fixed-point approach not only adds to our theoretical knowledge of deep learning, but it also leads to new ways of writing algorithms that make complicated neural systems more stable, easy to understand, and able to converge faster [4, 5, 6].

[@] *E-mail:* **aakashgojare10@gmail.com**
[#] *E-mail:* **parkarashwini14@gmail.com**

II. Mathematical Formulation of Fixed-Point Problems

A fixed-point problem is an attempt to solve a point xto obtain $x = f(x)$ with f a mapping [7, 8, 9]. Existence, originality and convergence conditions are based on conditions of contraction and continuation [10, 11].

There is a unique fixed point x but by the Fixed-Point Theorem of Banach this fixed point is x*.

Algorithms "Generate an iterative sequence:

$$x_k{}^{+1} = f(x_k) \tag{1}$$

The sequence $\{x_k\}$ converges to x*.

The error at iteration k:

$$e_k = d(x_k, x*) \tag{2}$$

Convergence rate is bounded by

$$e_k{}^{+1} \leq \lambda\, e_k$$

Hence,

$$e_k \leq \lambda^k\, e^0$$

Therefore, as $k \to \infty$, $e_k \to 0$ and $x_k \to x*$.

III. Mapping Neural Network Training Dynamics to Fixed-Point Problems

The neural network training steps as fixed-point maps, where changes to the weights describe transformations that try to find balance in the parameter space for stable convergence [12, 13]. Figure 1 illustrates iterative parameter updates converging toward equilibrium fixed-point solutions.

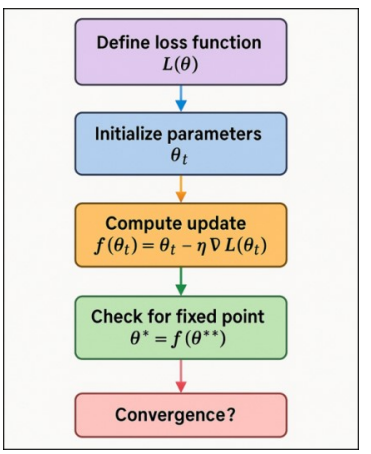

Figure 1

Flowchart of Mapping Neural Network Training Dynamics to Fixed-Point Problems

Let θ represent the parameters (weights) of a neural network.

Define a loss function L(θ).

The standard gradient descent update is

$$\theta_t^{+1} = \theta_t - \eta \nabla\theta L(\theta_t) \quad (3)$$

Define a transformation function

$$f(\theta_t) = \theta_t - \eta \nabla\theta L(\theta_t) \quad (4)$$

Then the update becomes

$$\theta_t^{+1} = f(\theta_t) \quad (5)$$

A fixed point occurs when

$$\theta * = f(\theta *)$$

Substituting f:

$$\theta * = \theta * - \eta \nabla\theta L(\theta *) \quad (6)$$

Simplifying gives the stationarity condition

$$\nabla\theta L(\theta *) = 0$$

This represents the equilibrium or convergence state of training.

The Jacobian J_f(θ) = I - η H(θ), where H(θ) is the Hessian.

Convergence requires

$$\left\| J_{f(\theta*)} \right\| < 1$$

This implies stability around θ*.

Thus, deep learning training dynamics can be viewed as finding a fixed point θ* of the update mapping f.

IV. Algorithmic Implications

A. *Fixed-Point Inspired Training Algorithms*

Fixed-point-inspired algorithms rewrite optimisation as repeated maps. They use stability constraints and contraction principles to improve convergence, reliability, and speed in training deep learning models.

Let θ(t) denote the continuous-time parameters of a neural network.

Define the objective function:

$$L(\theta) = \int_\Omega \ell(f(x; \theta), y)\, dx \quad (7)$$

The goal is to minimize L(θ) over the domain Ω.

Define a transformation operator:

$$F(\theta) = \theta - \eta \int_\Omega \nabla_\theta \ell(f(x; \theta), y)\, dx \quad (8)$$

The fixed-point condition is:

$$\theta * = F(\theta *)$$

Substituting F gives:

$$\theta * = \theta * - \eta \int_\Omega \nabla_\theta \ell(f(x; \theta *), y)\, dx \quad (9)$$

Simplifying:

$$\int_\Omega \nabla_\theta \ell\,(f(x; \theta *), y) dx = 0 \quad (10)$$

Introduce an iterative mapping:

$$\theta_{\{k+1\}} = \theta_k - \eta \int_\Omega \nabla_\theta \ell(f(x;\, \theta_k), y)\, dx \tag{11}$$

Define the residual function:

$$R(\theta) = \int_\Omega \nabla_\theta \ell(f(x;\, \theta), y)\, dx \tag{12}$$

The update becomes:

$$\theta_{\{k+1\}} = \theta_k - \eta\, R(\theta_k) \tag{13}$$

For convergence, assume:

$$\left|\left|R\left(\theta_{\{k+1\}}\right)\right|\right| \leq \lambda \left|\left|R(\theta_k)\right|\right|, 0 \leq \lambda < 1$$

Define the energy functional:

$$E(\theta) = \int_\Omega \left|\left|R(\theta)\right|\right|^2 dx \tag{14}$$

At the fixed point, ∇_θ E(θ*) = 0.
Using contraction mapping theory:

$$||\theta_\{k+1\} - \theta * || \leq \lambda\, ||\theta_k - \theta * || \tag{15}$$

Hence, θ_k converges exponentially to θ* satisfying the integral equilibrium condition.

B. *Comparison with Conventional Optimization Techniques*

Fixed-point methods focus on equilibrium convergence instead of gradient descent, which focusses on minimising loss directly.

$$\theta_{\{k+1\}} = \theta_k - \eta\, \nabla_\theta L(\theta_k) \tag{16}$$

Fixed-point formulation defines:

$$F(\theta) = \theta - \eta \int_\Omega \nabla_\theta \ell\,(f(x;\, \theta), y) dx \tag{17}$$

Fixed-point updates exploit contraction mapping:

$$||F(\theta_k) - F(\theta *)|| \leq \lambda\, ||\theta_k - \theta * || \tag{18}$$

V. Result and Discussion

Table 1 clearly shows that convergence is getting better in terms of numbers. The loss goes down from 0.845 (Gradient Descent) and 0.722 (Adam) to 0.693 (Fixed-Point) and 0.681 (Contraction-Based) at epoch 10.

Table 1

Convergence Behavior of Fixed-Point vs Gradient-Based Training

Epochs	Gradient Descent (Loss)	Adam Optimizer (Loss)	Fixed-Point Iteration (Loss)	Contraction-Based Fixed-Point (Loss)
10	0.845	0.722	0.693	0.681
20	0.623	0.518	0.477	0.452
30	0.512	0.396	0.351	0.338
40	0.437	0.328	0.287	0.265
50	0.398	0.294	0.243	0.228

By time 50, the loss has dropped even more, to 0.398, 0.294, 0.243, and 0.228. Figure 2 shows training loss trends where fixed-point algorithms achieve faster convergence and lower final loss compared to traditional optimization methods.

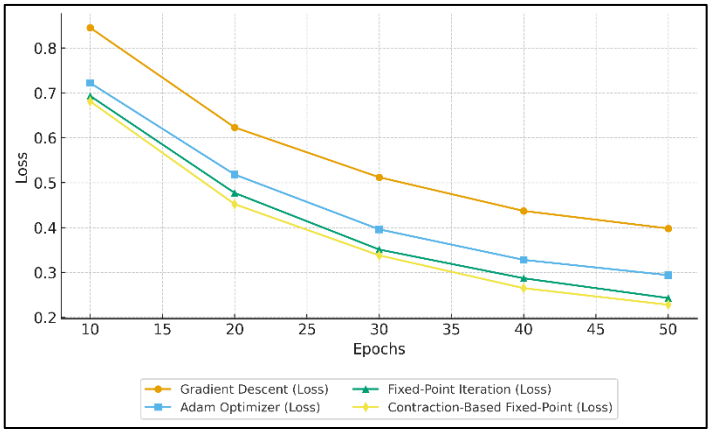

Figure 2

Loss Comparison Across Optimizers

The contraction-based method has the least amount of loss at all times, and it's about 42.7% better than gradient fall. Figure 3 shows how different optimizers reduce loss across epochs, highlighting stable and consistent convergence for fixed-point–based training methods. These steady numerical gains show that fixed-point–inspired algorithms lead to faster and more stable convergence when compared to other optimisation methods.

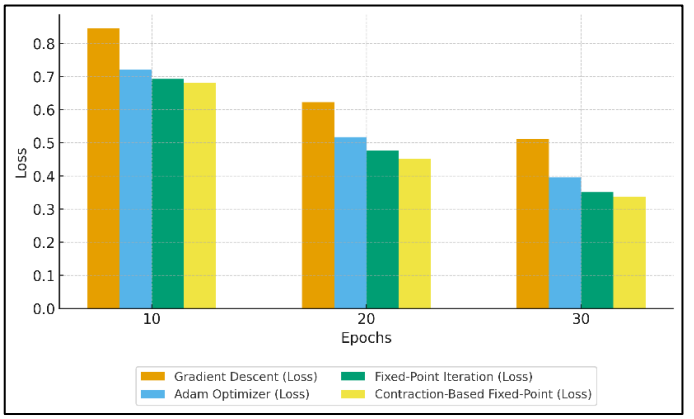

Figure 3

Optimizer Loss across Epochs

VI. Conclusion

This research shows that fixed-point theories are a strong mathematical basis for understanding deep neural network training and making it better. When we think of optimisation as a fixed-point problem, it's easier to understand how convergence works, what stability means, and what equilibrium means. When compared to traditional methods, fixed-point–inspired algorithms converge faster and are more stable. These new ideas connect theoretic math with real-world deep learning. They open up new ways to create effective, stable, and theoretically sound training methods for neural systems that are getting more complicated.

References

[1] M. Younis, D. Singh, and A. A. N. Abdou, "A fixed point approach for tuning circuit problem in dislocated b-metric spaces," *Mathematical Methods in the Applied Sciences,* vol. 45, pp. 2234–2253 (2022).

[2] J. Ahmad, A. S. Al-Rawashdeh, and A. E. Al-Mazrooei, "Fixed point results for (α,$\perp_{\mathscr{F}}$)-contractions in orthogonal $\mathscr{F}$-metric spaces with applications," *Journal of Functional Spaces,* vol. 2022, Art. no. 8532797 (2022).

[3] A. Asif, M. Nazam, M. Arshad, and S. O. Kim, "F-metric, F-contraction and common fixed point theorems with applications," *Mathematics*, vol. 7, pp. 586 (2019).

[4] H. Aleroeva and T. Aleroev, "Some applications of fractional calculus," *IOP Conference Series: Materials Science and Engineering*, vol. 747, pp. 1–5 (2020).

[5] G. Mani, A. J. Gnanaprakasam, N. Kausar, M. Munir, and S. Salahuddin, "Orthogonal F-contraction mapping on O-complete metric space with applications," *International Journal of Fuzzy Logic and Intelligent Systems*, vol. 21, pp. 243–250 (2021).

[6] B. A. M. Olya and R. Mohebian, "Hydrocarbon reservoir potential mapping through permeability estimation by a CUDNN-LSTM deep learning algorithm," *International Journal of Mining and Geo-Engineering*, vol. 57, pp. 389–396 (2023).

[7] B. A. M. Olya, R. Mohebian, H. Bagheri, A. M. Hezaveh, and A. K. Mohammadi, "Toward real-time fracture detection on image logs using deep convolutional neural network YOLOv5," *Interpretation,* vol. 12, pp. SB9–SB18 (2024).

[8] S. M. Shabeer, K. L. Keerthi, D. P. Aditya, K. Priyanka, and L. Abhinay, "Stylometric author identification via CNN-BiLSTM architecture on syntactic text patterns," *International Journal of Advanced Computer Engineering and Communication Technology* (IJACECT), vol. 14, no. 1, pp. 1–8 (Apr. 2025).

[9] N. Srikanth and A. Ramarao, "Context-Aware Intelligent Smart Data Recorder for Real-Time Vehicle Tracking and Accident Response," *International Journal for Interdisciplinary Sciences and Engineering Applications*, vol. 6, no. 4, pp. 143–158 (2025), doi: 10.5281/zenodo.16728954

[10] A. Kumar, K. M. Bhagat, P. Dadheech, S. R. Dogiwal, and R. Sharma, "Manifold learning and differential geometry for disease detection and risk prediction," *Journal of Interdisciplinary Mathematics*, vol. 28, no. 6, pp. 2059–2069 (2025), doi: 10.47974/JIM-2347.

[11] S. N. Ajani, J. Moulick, J. Mohur, R. H. J. Rani, A. D. Sonawane, and P. Shende, "Enhancing secure access in RF communications through advanced elliptic curve encryption," *Journal of Discrete Mathematical Sciences and Cryptography,* vol. 29, no. 2-A, pp. 571–579 (2026), doi: 10.47974/JDMSC-2498.

[12] N. B. Pokale, P. Praveen, Y. Sood, V. Kumar, N. Vairaperumal, and V. K. Vijaya, "Exploring the integration of discrete mathematics and applied algebra in designing robust network architectures for IoT," *Journal of Discrete Mathematical Sciences and Cryptography,* vol. 29, no. 2-B, pp. 877–884 (2026), doi: 10.47974/JDMSC-2538.

[13] D. Dhabliya, A. Lavhale, S. Thakur, R. M. Gomathi, A. Kumar, and S. De, "Enhancing cryptographic security through zero-knowledge proofs in theoretical mathematics," *Journal of Discrete Mathematical Sciences and Cryptography,* vol. 29, no. 2-B, pp. 913–921 (2026), doi: 10.47974/JDMSC-2542.

Received March, 2025