

On the use of evolutionary and swarm intelligence algorithms in transfer learning approaches: a review

Abstract

Transfer Learning (TL) utilizes pre-trained models to solve similar problems. The knowledge from the original model is transferred to a new model during training, aiming to leverage previous knowledge in a new task. Natural Computing (NC) algorithms, such as Evolutionary Computation (EC) and Swarm Intelligence (SI), draw inspiration from nature, adapting more easily to new computational problems. This bio-inspired adaptation can enhance the performance of TL techniques, improving generalization and reducing computational costs. We investigate how evolutionary and swarm-intelligence algorithms are applied in TL, their contributions, the addressed problems, and the conducted experiments. We employ a systematic review following the PRISMA protocol, PICOS strategy, and START software to analyze primary studies.

Keywords: systematic review, protocol, algorithm, swarm intelligence, transfer learning, natural computing

Volume 8 Issue 2 - 2023

Rita Xavier,¹ Leandro Nunes de Castro^{1,2,3}

¹School of Technology, Universidade Estadual de Campinas (UNICAMP), Brazil

²Department of Computing & Software Engineering, Florida Gulf Coast University (FGCU), USA

³In.lab iNovaHC, Clinics Hospital, School of Medicine, University of São Paulo (USP), Brazil

Correspondence: Leandro Nunes de Castro, School of Technology, Universidade Estadual de Campinas (UNICAMP), Florida Gulf Coast University (FGCU), 10501 FGCU Blvd. South, Fort Myers, Florida, USA, Tel +12395901362, Email leandron@unicamp.br, ldecastrosilva@fgcu.edu

Received: December 10, 2023 | **Published:** December 26, 2023

Abbreviations: TL, Transfer Learning; NC, Natural Computing; NIC, Nature-Inspired Computing; EC, Evolutionary Computation; SI, Swarm Intelligence; SR, Systematic Review.

Introduction

Transfer Learning (TL) is an approach in which a previously trained model shares its effective characteristics with other models to solve or optimize the solution to similar problems. The knowledge acquired by this pre-trained model, known as the *source domain*, which has demonstrated satisfactory performance in optimizing a prior task, called the *source task*, is transferred to a new model in training, called the *target domain*. This new model may be similar to or different from the previous one. The goal is to train the new model to solve a new task, called the *target task*, which may or may not be like the source task.¹ Natural Computing (NC) studies and models natural processes, providing alternative solutions to problems not efficiently addressed by traditional techniques.

Nature-inspired algorithms, such as Evolutionary Computation (EC) and Swarm Intelligence (SI), illustrate how the behavior of living organisms can inspire technological advances.² Given the inspiration from nature, NC algorithms can adapt to the environment and time. For example, Genetic Algorithms (GA) can undergo genetic variations when creating new populations, while Particle Swarm Optimization (PSO) adjusts its behavior based on the swarm to which it belongs.³ The adaptation of these algorithms, including their learning capability, is a relevant area of study in computer science.

Thus, it is possible to expect that applying the capabilities of nature-inspired algorithms can improve the performance of TL techniques, bringing greater generalization and even lower computational cost to the knowledge transfer iterations performed within them. Therefore, it is essential to understand how these bio-inspired algorithms are being applied in TL techniques, how they assist in this application, what problems (both in TL and NC) are being solved or optimized, and how experiments were conducted (including instances, benchmarks, and performance evaluation metrics) to compare the proposed algorithms with state-of-the-art algorithms.

This approach was conducted through a systematic review, a methodology that employs robust and reproducible protocols to address questions formulated by the authors. In this study, we adopted the PRISMA protocol,⁴ the PICOS strategy,⁵ and the START software⁶ to systematize the collection of works and analyze each of the primary studies.

The paper is organized as follows. Section 1 provides the context, motivation and a general overview of the work. Section 2 presents the main concepts explored in the paper, emphasizing transfer learning, evolutionary computing, and swarm intelligence. Section 3 presents the materials and methods used in the research, and Section 4 presents the results and discussions around them. The paper is concluded in Section 5 with a general overview of its contributions and the main open problems observed in the field.

Theoretical background

In this review, certain concepts are discussed that need clarification for a better understanding of the content. Therefore, the following subsections present a general overview of nature-inspired algorithms and transfer learning.

A briefing on nature-inspired computing

Natural Computing (NC) is a branch of computer science that seeks inspiration from nature to create innovative approaches in solving complex problems, exploring naturally intelligent systems. This interdisciplinary field combines computer science with disciplines such as chemistry, physics, and biology.²

Some authors^{2,4} classify the terminology of Natural Computing into three main categories:

1) Simulation of Natural Phenomena in Computing; 2) Nature-Inspired Computing (NIC), which develops techniques based on nature to solve problems; and 3) Computing with New Materials from Nature, using materials collected from nature to compute.⁷

Figure 1 illustrates the main areas and subareas of Natural Computing.

There is a trend in Natural Computing (NC) research, especially in the category of Nature-Inspired Computing (NIC), where the emphasis is given to Artificial Neural Networks (ANN). These networks play a significant role in emerging machine learning methodologies, such as Deep Learning,⁸ and in other technologies, such as Natural Language Processing (NLP).⁹ Therefore, investigations dedicated to exploring other categories of NC can identify gaps and introduce new paradigms to the field of study.

The systematic review conducted in this work provides an overview of NC algorithms that can assist in Transfer Learning techniques. The algorithms proposed in the base studies are mostly in the Nature-Inspired Computing (NIC) class, specifically in the categories of Evolutionary Computing (EC) and Swarm Intelligence (SI). In the following sections, we will explore the fundamental concepts of these algorithms.

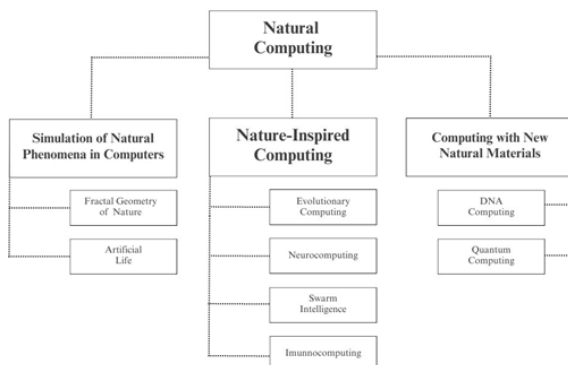


Figure 1 Natural computing categories.

Evolutionary Computing (EC)

During this review, it was observed that most NC algorithms used in TL techniques fall within the context of evolutionary computation. These algorithms, based on evolutionary biology, aim to adapt populations of individuals to the environment through reproduction, genetic variation, and selection.³ These techniques involve reproduction, random variation, and selection, fundamental processes of evolution.¹⁰ Evolution, as an optimization process, aims for adaptation to the environment, as highlighted by Mayr¹¹ and related to computation by Baeck¹⁰. The ability for robust adaptation, essential for computational systems, results in agile and precise performance, defining an algorithm as optimal.¹²

Among the algorithms observed in the primary studies, there are two subtypes of EC algorithms commonly used with TL strategies, namely Genetic Algorithms (GA) and Genetic Programming (GP). Each category of algorithms in evolutionary computation has distinct characteristics. GAs are based on search, often using either a binary or a real-valued representation of solutions, with a tendency toward representation problems due to their generality, replicating recombination and mutation processes across generations.¹³ On the other hand, GPs deal with representation problems, increasing the adaptability of structures,¹⁴ using hierarchies composed of terminal and primitive functions.¹⁵

Swarm Intelligence (SI)

Another Nature-Inspired Computing approach observed in the base studies is that of Swarm Intelligence (SI). SI occurs when a population of limited and usually non-intelligent entities exhibits swarm behavior, interacting intelligently with each other,¹⁶ as observed in flocks of birds, herds of quadrupeds, or schools of fish. This concept extends to various other 'populations', such as ant colonies, cars in traffic, or even people in a crowd, characterized by the collective behavior of the group.¹⁷

There are two lines of research in Swarm Intelligence:³ 1) Studies based on insects or animals with social behaviors; and 2) Works that explore the ability of human society to process knowledge. Although the studied individuals are distinct, both depend on interaction within a population and with the surrounding environment, resulting in environmental impacts or interactions between individuals.

The SI approach present in the base studies is Particle Swarm Optimization (PSO), which simulates the social behavior of a swarm. In this type of algorithm, particles use their best positions and those of their neighbors to seek an optimal solution.² Individuals learn from their own experiences and those of others, evaluating and imitating the behavior of the most successful. PSO is motivated by the behavior of fish schools and bird flocks, seeking a balance between global and local optimization through a self-organized and self-adjusting system.¹⁷

A briefing on transfer learning

Transfer Learning (TL) optimizes the application of knowledge across domains, potentially saving computational resources, improving training efficiency, promoting better generalization, and facilitating new task learning. This approach allows models to inherit knowledge, resulting in better performance, faster convergence, and greater adaptability to limited data. Some authors have interesting and comprehensive studies about TL.^{1,18,19} TL operates on the premise that knowledge obtained in one context can be applied in another, optimizing the learning process. This approach understands the idea that, just as humans build upon their existing knowledge when learning new skills or concepts, learning models should have the ability to do the same. The essence is to efficiently leverage acquired knowledge, training models to adapt and generalize to new tasks or domains.

The TL technique is characterized by adapting a model or system from a source domain, denoted as SD, with a corresponding source task, ST, to improve the performance of a target task, TT, in a target domain, TD, where $SD \neq TD$ and $ST \neq TT$. Formally, TL can be represented as follows:

$$SD, ST \rightarrow TD, TT$$

In this scenario, \rightarrow represents the knowledge transfer process, whose main goal is to facilitate the learning of TT, capitalizing on insights and experiences gained from ST.

Transfer Learning Strategies

Transfer Learning (TL) can be categorized into four configurations of transfer strategies:¹

- I. Instance Transfer:** Addresses differences in distribution margins using reweighted instances.²⁰
- II. Feature Representation Transfer:** Mitigates dissimilarities in domain feature spaces, which can be either symmetric or asymmetric.^{21,22}
- III. Parameter Transfer:** Propagates knowledge through shared parameters between learner models from different domains.²³
- IV. Relational Knowledge Transfer:** Involves acquiring shared relationships between domains, facilitating the construction of a shared knowledge mapping.²⁴

The other categorization of TL is based on the learning strategy and can be divided into three main concepts:¹⁹

- I. Inductive Learning:** Applies knowledge from a previous task to a new task with distinct objectives.

- II. Transductive Learning:** Maintains tasks but changes domains, adapting the model for better performance.
- III. Unsupervised Learning:** Transfers knowledge from tasks with unlabeled data to tasks with scarcely labeled data, alleviating the need for annotation in challenging or computationally expensive scenarios.

Materials and methods

This study aims to explore how evolutionary and swarm algorithms are used in Transfer Learning techniques, identifying the objectives of this usage and evaluating the performance of these approaches. Following the 2020 PRISMA Guidelines,⁴ the PICOS strategy⁵ and the START software,⁶ a systematic review protocol was developed to analyze primary studies, aiding in understanding the objectives of this paper. This review addresses the following research questions:

- I. Question 1:** Which natural computing algorithms are being used to optimize Transfer Learning techniques?
- II. Question 2:** What Transfer Learning problem is being optimized by the natural computing algorithm?
- III. Question 3:** Which Transfer Learning technique is the target of the natural computing algorithm, how is the performance evaluated, which instances were used, and what were the comparison algorithms (benchmarks)?

To achieve our goal, search strings consistent with the primary studies sought were created, along with the OR, NOT, AND operators, or their equivalents (-, +, =) applied to search engines. The used strings were:

- I. Query 1:** “transfer learning” AND “swarm intelligence” NOT “deep learning”
- II. Query 2:** “transfer learning” AND “evolutionary algorithm” NOT “deep learning”
- III. Query 3:** “transfer learning” AND “evolutionary computing” NOT “deep learning”.

The reason why we excluded the deep learning papers is because they open avenue for a completely different line of research and would result in a review that does not cover the use of evolutionary and swarm algorithms to improve transfer learning. Instead, a plethora of works involving the use of TL in deep learning approaches would be returned, but these are reviewed in other works from the literature.^{8,9}

After searching and collecting data from the databases, some exclusion criteria were applied for better utilization of the primary studies. We used the Scopus, IEEE Xplore, and Web of Science databases with a publication date range from 2018 to 2023. Articles were excluded if they met the following exclusion criteria:

- I. Not peer-reviewed and/or in a pre-publication stage.
- II. Conducted before the year 2018.
- III. In languages unknown to the authors.
- IV. Do not address the application of a nature-inspired algorithm in transfer learning techniques.
- V. Focus on transfer learning applications in deep learning.
- VI. Do not propose a new method or algorithm.

After primary extraction with the temporal criterion applied, 2,723 studies were retrieved (Table 1). Based on the inclusion and exclusion criteria, these works were reviewed (title, authorship, year, abstract, and keywords), and a total of 60 studies remained for full-text

analysis. After reading and analyzing these texts, 10 were included in the systematic review.

Table 1 Primary outcomes

Query	IEEE	Scopus	Web of Science
1	25	446	96
2	218	1.233	284
3	63	115	243

Results and discussions

The objective of this systematic review is to understand the use of natural computing algorithms for the benefit and improvement of TL techniques. To address all research questions in this study, the paper will be divided into three main topics and their subtopics.

- I. Transfer Learning, Evolutionary, and Swarm Algorithms:** A summarization of observations during this review will be presented to address the first research question.
- II. Objectives of Evolutionary and Swarm Algorithms in Transfer Learning Techniques:** This topic will address the objectives of applying evolutionary and swarm algorithms in the proposed algorithms or methods.
- III. Experiments with Proposed Methods and Algorithms:** The third topic will present how experiments with the proposed methods or algorithms were conducted, indicating performance measures, comparative algorithms, and instances used.

Transfer learning, evolutionary and swarm algorithms

The first question formulated for this review was “which natural computing algorithms are being used to optimize transfer learning techniques?” To understand this question, it is necessary to observe how NC algorithms can assist in the performance of TL techniques.

Two categories of NC algorithms were identified among the analyzed primary studies, namely SI and EC, specifically PSO, GA, and GP. The algorithms proposed by the analyzed primary studies are applying nature-inspired algorithms to solve a variety of problems, both in optimization in NC and in optimizing TL techniques themselves. Table 2 shows the distribution and summarization of these studies.

Table 2 Review summarization

Transfer Learning Technique	Nature-Inspired Algorithm	References
	Genetic Algorithm	Alghamdi et al. ²⁵
Instance Weighting	Genetic Programming	Chen et al. ²⁶
Feature Selection-Based Domain Adaptation	Particle Swarm Optimization	Nguyen et al. ²⁷ Sanodiya et al. ²⁸ Castillo-García et al. ²⁹ Dhrif et al. ³⁰
Feature Transformation Based Domain Adaptation	Genetic Programming	Al-Helali et al. ³¹ Al-Helal et al. ³² Al-Helali et al. ³³
Parameter-Selection	Particle Swarm Optimization	Sanodiya et al. ³⁴

The most common TL techniques among the primary studies collected for this review are “Instance Weighting,” “Feature Selection-Based Domain Adaptation,” “Feature Transformation-Based Domain Adaptation,” and “Parameter Selection”.

In Instance Weighting, three algorithms proposed in the base studies are Genetic Algorithm for Source Task Problems – GA(S),²⁵ Genetic Algorithm for Target Task Problems – GA(T),²⁵ and Instance Transfer Genetic Programming – ITGP.²⁶ Instance Weighting, a statistical bias correction technique, addresses sample selection bias, playing a crucial role in transfer learning to overcome the disparity between source and target domains. It reassesses data from the source domain, discarding harmful instances and reusing relevant ones in the target domain, correcting the distribution difference between them.³⁵

It is understood that TL involves two domains, the source and the target. When they share the same feature space, i.e., $SD = TD$, transfer learning is referred to as Domain Adaptation. Most feature-based domain adaptation strategies aim to create a latent feature space acting as a connection between diverse domains. In Feature Selection-Based Domain Adaptation, the expectation is that datasets derived from two original sets, when projected into the new feature space, exhibit greater proximity than the original sets.²⁷

Among the base studies, five algorithms are identified: Semi-Supervised Particle Swarm Optimization (SemPSO),²⁷ Unsupervised Particle Swarm Optimization (UnPSO),²⁷ Feature Selection-Based Transfer Learning Approach Using Particle Swarm Optimization for Unsupervised Transfer Learning (FSUTL-PSO),²⁸ Sticky Binary Particle Swarm Optimization (SBPSO), and Feature Selection Algorithm for Transfer Learning (COMB-PSO-TL).³⁰

In addition to feature-based adaptation, there is Feature Transformation-Based Domain Adaptation, or Feature Weighting. In this TL technique, two approaches exist: symmetric and asymmetric.³⁶ In the symmetric approach, feature transformation maps both domains to a third one representing the common feature space, created to unify both domains. In contrast, asymmetric approaches aim to map one domain into the other.^{31–33} Three algorithms are associated with this technique: Multi-Tree Genetic Programming-Based Transfer Learning (MTGPTL),³¹ Multi-Tree Genetic Programming Domain Adaptation (MTGPDA),³² and Multi-Tree Genetic Programming-Based Transfer Learning (MTGP-Based TL).³³

As for the Parameter Selection technique, learning approaches transfer knowledge at the model/parameter level. For instance, when categorizing objects, knowledge of categories in the source domain can be transferred to the target domain using attributes such as shape and color.³⁷ These parameters, reflecting the probabilistic distribution of features, are learned from the source domain to facilitate classifier training in the target domain, enabling knowledge transfer at parameter level.¹⁹ For the Parameter Selection technique, there is only one algorithm, the Particle Swarm Optimization-based Parameter Selection Approach for Unsupervised Discriminant Analysis (UDATL-PSO).³⁴

Objectives of evolutionary and swarm algorithms in transfer learning techniques

The second question raised by this review pertains to understanding the objectives, that is, the problems being addressed within Transfer Learning techniques by the evolutionary and swarm algorithms. In the context of natural computing, these objectives are diverse, bringing forth interesting strategies to assist in potential challenges present in Transfer Learning techniques. Therefore, the following sections will describe which of these strategies were employed in the algorithms proposed in the foundational studies.

Evolutionary Computing Algorithms

Among the evolutionary computing algorithms discussed in the foundational studies, there are two approaches that use Genetic Algorithms (GA)²⁵ to assist in Transfer Learning techniques, and four

that use Genetic Programming (GP).^{26,31–33} Both GA(S) and GA(T)²⁵ apply GA to model the TL process as a search problem, allowing them to control and measure various factors affecting transfer performance, such as problem difficulty, problem similarity, and information transfer methods. The goal of GA(S) and GA(T) is to understand the underlying mechanisms and principles of TL, aiming to develop more effective knowledge transfer techniques.

With Genetic Programming algorithms, we have Instance Transfer Genetic Programming (ITGP).²⁶ ITGP seeks optimal weights for instances in the source domain, enabling the efficient reuse of informative instances from the source domain while mitigating the impact of harmful instances from the source domain. In summary, ITGP utilizes a new instance weighting framework to guide the evolutionary process through the evolution of weight vectors

In the Multi-Tree Genetic Programming-Based Transfer Learning (MTGPTL)³¹ algorithm, Genetic Programming aims to perform a feature-based transformation on the source domain data. The transformed data is subsequently used to train a model. On the other hand, Multi-Tree GP Domain Adaptation (MTGPDA)³² aims to construct a transformation that maps the feature space of the source domain to the feature space of the target domain. The last proposed algorithm based on GP is Multitree Genetic Programming-Based Transfer Learning (MTGP-Based TL),³³ which aims to construct a weighting that maps a source domain to a different but related target domain. This mapping is achieved using MGPT (Multitree Genetic Programming), constructing multiple features from the source domain's features so that the transformed data has a representation like that of the target domain. This results in more effective learning for the target domain, compensating for the lack of knowledge in the target domain due to the absence and scarcity of instances.

Swarm Intelligence Algorithms

In the realm of Swarm Intelligence (SI) algorithms, there are six representatives, all falling under the domain of Particle Swarm Optimization (PSO). In one of the basic studies,²⁷ two algorithms, SemPSO and UnPSO, are introduced. Despite employing different transfer strategies (one using transductive strategy and the other using unsupervised strategy, as their names suggest), the common goal of utilizing PSO in both algorithms is to propose a new fitness function that allows the PSO-based feature selection algorithm to operate when class label information is available (semi-supervised) and when it is not available in the target domain (unsupervised). In other words, the task of PSO is to search for feature subsets with the lowest fitness value.

In the Feature Selection-Based Transfer Learning Approach Using Particle Swarm Optimization for Unsupervised Transfer Learning (FSUTL-PSO)²⁸ algorithm, the nature-inspired algorithm aims to select a good subset of features between the source and target domains to eliminate the risk of degenerated feature transformation. PSO constructs a common feature space using features with high invariant characteristics. The Sticky Binary Particle Swarm Optimization (SBPSO)²⁹ utilizes PSO as a strategy to achieve an effective representation of invariant features that preserves important information from the original features.

The Feature Selection Algorithm for Transfer Learning (COMB-PSO-TL)³⁰ employs PSO to provide high discriminative capacity in both domains, minimize the difference between conditional distributions, and minimize the difference between marginal distributions in the source and target domains. In other words, it enhances classification accuracy in a target domain through knowledge transfer from a source domain. The algorithm conducts subset selection of features (genes) for transfer learning using a two-level combinatorial particle swarm

optimization algorithm. Finally, the Particle Swarm Optimization-based Parameter Selection Approach for Unsupervised Discriminant Analysis (UDATL-PSO)³⁴ uses PSO to select the best parameter values for knowledge transfer.

Experiments with the proposed methods and algorithms

The third question addressed in the research pertains to the

problems and applications that the proposed methods aim to optimize. In other words, by the end of this section, we aim to comprehend the problems tackled by the proposed algorithms and how comparative experiments (benchmarks) were conducted. To achieve this, Table 3 summarizes the type of optimization, the problems, and the names of the proposed algorithms, the instances used, the benchmark algorithms presented, and the performance metrics applied for conducting these comparisons.

Table 3 Experiments information & proposed algorithms and methods

TL Problem	NC Problem	Proposed algorithm	Performance measures	Instances	Bench markings
Instance Weighting	Function Approximation	GA(S) ²⁵	Diversity	Static Source Problem Fixed Target Problems	Estimation Distance (ED) 30-Top, 30-Best 100-Top, and 100-Best
		GA(T) ²⁵	Generations for Convergence Friedman Dataset		
	RSE Median	UCI and Delve Datasets (Wine, Kin, Student, House, Abalone)	GP-Tar GP-Comb TLGP		
Feature Selection-Based Domain Adaptation	Symbolic Regression	ITGP ²⁶	Significance Wilcoxon Test Z-Test	Friedman- I Dataset Gas Sensor Dataset	TLGP-NS SVR-Tar SVR-W
		SemPSO ²⁷ UnPSO ²⁷	Accuracy Significance Wilcoxon Test	Handwritten Digits Datasets Objects Recognition Dataset	TCA and STCA MIDA and SMIDA
	Classification	FSUTL-PSO ²⁸	Accuracy Parameter Sensitivity Tests	Office+Caltech Dataset PIE Face Recognition Dataset	TCA, TSL, TJM, TDA-AL mSDA, MEDA RDALR, RTML GFK, GA-DA AS, UnPSO JDA, JGSA LTSL, ILS CMDL, CORAL, CDDA DGA-DA, BDA
		SBPSO ²⁹	Accuracy Precision Recall F1-Score Area Under the Curve (AUC)	Gas Sensor Drift Dataset Handwritten Digits Dataset Prostate Dataset TripAdvisor Dataset Source Synthetic Dataset	DAN TCA ATL Simple PSO COMB- PSO-LS COMB- PSO-TF
Feature Transformation	Symbolic Regression	COMB-PSO-TL ³⁰	Accuracy Mean	Target Synthetic Dataset UCI Datasets (Housing,	
		MTGPPTL ³¹	Standard Deviation RSE	Concrete, Forestfires, Yacht-Hydrodynamics)	NoTL PlainTL
		MTGPDA ³²	RSE	UCI Datasets (Housing, Concrete, Yacht, Forestfires, Auto- MPG, Imports)	WKNNIM WKNNIMDC RFIM MTGPPTL
Parameter- Selection	Classification	MTGP-Based TL ³³	RSE	UCI Datasets (Housing, Concrete, Yacht, Forestfires, Auto- MPG, Imports)	Traditional TL Methods
		UDATL-PSO ³⁴	Accuracy Time/Space Complexity	PIE Face Recognition Caltech Office Datasets	Non-TL Algorithms TL Algorithms

These algorithms are proposed not only to optimize challenges encountered in Transfer Learning (TL) techniques but also to assist in broader optimization problems such as Function Approximation, Classification, and Symbolic Regression. These general optimization problems are present to conduct comparative tests between the proposed algorithms and benchmark algorithms—state-of-the-art algorithms that also engage in some form of optimization. The Function Approximation problem aims to find a function that best fits a set of observed data or represents a desired relationship between variables. In other words, the goal is to find a mathematical expression that is a good approximation for the real function relating to the variables of interest. The GA(S) and GA(T)²⁵ algorithms aim to test TL techniques in frameworks based on Genetic Algorithms. In this experiment, the authors created an environment with source and target problems to test various estimations, such as distances and population numbers generated by the methods.

In the Classification problems, the objective is to assign a category or label to a given object based on its observed characteristics.³⁸ Assigned to this problem are the algorithms SemPSO,²⁷ UnPSO,²⁷ FSUTL-PSO,²⁸ SBPSO,²⁹ and COMB-PSO-TL.³⁰ These algorithms are applied to image and document classification using datasets such as Office+Caltech,³⁹ PIE Face Recognition,⁴⁰ Gas Sensor,⁴¹ Handwritten Digits,^{42,43} Prostate,⁴⁴ and TripAdvisor.⁴⁵ For Regression problems, the goal is predicting or estimating a continuous value based on observed characteristics.²⁸ The primary studies analyzed involve Symbolic Regression, which aims to discover a symbolic expression capturing the underlying structure of the data, providing interpretability to the model. This is particularly useful for missing values, a problem explored by the three Genetic Programming-based algorithms (MTGPTL,³¹ MTGPDA,³² and MTGP-Based TL³³).

In summary, the algorithms addressed in this study not only aim to optimize specific challenges encountered in TL techniques but also extend to assisting in broader optimization problems. By addressing these general problems, the algorithms undergo comparative tests, facing benchmark algorithms representing the state-of-the-art in optimization. Collectively, these approaches open valuable perspectives for advancements in both Transfer Learning and more comprehensive optimization problems.

Conclusions and future trends

This paper provided a comprehensive investigation into the crucial role of evolutionary and swarm algorithms in optimizing Transfer Learning techniques. By addressing three main questions, it was possible to explore not only the diversity of the algorithms applied in TL but also their specific objectives and the results of practical experiments. The importance of understanding these algorithms was highlighted, categorizing them into Swarm Intelligence (SI) and Evolutionary Computing (EC). The analysis revealed the significant presence of algorithms such as Particle Swarm Optimization (PSO), Genetic Algorithms (GA), and Genetic Programming (GP), showing the versatility of these approaches in overcoming the unique challenges encountered in TL. Innovative strategies employed by these algorithms to solve specific TL problems were explored. From searching for optimal weights for instances in the source domain to constructing feature-based transformations, the potential of these approaches to enhance the effectiveness of learning in TL became evident. The experiments conducted covered problems such as Function Approximation, Classification, and Symbolic Regression. The positive results of these experiments provided valuable insights into the practical applicability of these algorithms in real-world scenarios.

Considering these reflections, it becomes evident that natural computing algorithms play an essential role in the evolution of TL techniques. Their contributions not only optimize specific challenges in TL but also extend to broader optimization problems such as Function Approximation and Classification.

As future perspectives, further exploration of these techniques in more complex domains and adaptation to specific challenges in diverse areas is suggested. The continuous refinement of these algorithms is essential to tackle emerging challenges and maintain the relevance of these approaches at the forefront of artificial intelligence and machine learning.

Acknowledgments

This research is funded by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001, and by Grant #2021/11905-0, São Paulo Research Foundation (FAPESP).

Conflicts of interest

The authors declare there is no conflict of interest.

References

- Pan SJ, Yang Q. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*. 2009;22(10):1345–1359.
- Siddique N, Adeli H. Nature inspired computing: an overview and some future directions. *Cognitive computation*. 2015;7(6):706–714.
- De Castro LN. Fundamentals of natural computing: an overview. *Physics of Life Reviews*. 2007;4(1):1–36.
- Page MJ, Joanne E, McKenzie, Patrick M, Bossuyt, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *International journal of surgery*. 2020;88.
- Fabbri S, Cleiton Silva, Elis Hernandes, et al. *Improvements in the StArt tool to better support the systematic review process*. Proceedings of the 20th international conference on evaluation and assessment in software engineering. 2016;1–5.
- Amir-Behghadami M, Janati A. Population, Intervention, Comparison, Outcomes and Study (PICOS) design as a framework to formulate eligibility criteria in systematic reviews. *Emergency Medicine Journal*. 2020;37(6):387.
- De Castro LN. *Fundamentals of natural computing: basic concepts, algorithms, and applications*. CRC Press. 2006.
- Bharti V, Biswas B, Shukla KK. Recent trends in nature inspired computation with applications to deep learning. *10th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*. 2020;294–299.
- Young T, Hazarika D, Poria S, et al. Recent trends in deep learning based natural language processing. *IEEE Computational Intelligence Magazine*. 2018;13(3):55–75.
- Baek T, Fogel DB, Michalewicz Z. *Evolutionary computation 1: Basic algorithms and operators*. CRC press. 2018.
- Mayr E. The why and how of species. *Biology and Philosophy*. 1988;3:431–441.
- De Lacerda EG, De Carvalho ACPLF. Introdução aos algoritmos genéticos. *Sistemas inteligentes: aplicações a recursos hídricos e ciências ambientais*. 1999;99–148.
- Lambora A, Gupta K, Chopra K. *Genetic algorithm-A literature review*. 2019 international conference on machine learning, big data, cloud and parallel computing (COMITCon). 2019;380–384.

14. Koza JR. Genetic programming as a means for programming computers by natural selection. *Statistics and computing*. 1994;4:87–112.
15. Koza JR. *Survey of genetic algorithms and genetic programming*. Wescon conference record. 1995;589–594.
16. White T, Pagurek B. *Towards multi-swarm problem solving in networks*. Proceedings International Conference on Multi Agent Systems. 1998;333–340.
17. Eberhart RC, Shi Y, Kennedy J. *Swarm intelligence*. Elsevier. 2001:1–8.
18. Weiss K, Khoshgoftaar TM, Wang D. A survey of transfer learning. *Journal of Big data*. 2021;3(1):1–40.
19. Zhuang F, Qi Z, Duan K, et al. A comprehensive survey on transfer learning. *Proceedings of the IEEE*. 2020;109(1):43–76.
20. Huang J, Gretton A, Borgwardt K, et al. *Correcting sample selection bias by unlabeled data*. *Advances in neural information processing systems*. 2006;19.
21. Long M, Wang J, Ding G, et al. *Transfer feature learning with joint distribution adaptation*. Proceedings of the IEEE international conference on computer vision. 2013;2200–2207.
22. Feuz KD, Cook DJ. Transfer learning across feature-rich heterogeneous feature spaces via feature-space remapping (FSR). *ACM transactions on intelligent systems and technology (TIST)*. 2015;6(1):3.
23. Lawrence ND, Platt JC. *Learning to learn with the informative vector machine*. Proceedings of the twenty-first international conference on Machine learning. 2004. p. 65.
24. Mihalkova L, Mooney RJ. *Transfer learning by mapping with minimal target data*. In Proceedings of the AAAI-08 workshop on transfer learning for complex tasks. 2008;31–36.
25. Alghamdi T, Heckendorn RB. *An evolutionary computation based model for testing transfer learning strategies*. IEEE Congress on Evolutionary Computation (CEC). 2021;1380–1389.
26. Chen Q, Xue B, Zhang M. Genetic programming for instance transfer learning in symbolic regression. *IEEE Transactions on Cybernetics*. 2022;52(1):25–38.
27. Nguyen BH, Xue B, Andraea P. *A particle swarm optimization based feature selection approach to transfer learning in classification*. Proceedings of the genetic and evolutionary computation conference. 2018;37–44.
28. Sanodiya RK, Tiwari M, Mathew J, et al. A particle swarm optimization-based feature selection for unsupervised transfer learning. *Soft Computing*. 2020;24:18713–18731.
29. Castillo García G, Morán Fernández L, Bolón Canedo V. Feature selection for domain adaptation using complexity measures and swarm intelligence. *Neurocomputing*. 2023;548:126422.
30. Dhrif H, Bolón Canedo V, Wuchty S. *Gene Subset Selection for Transfer Learning using Bilevel Particle Swarm Optimization*. 19th IEEE International Conference on Machine Learning and Applications (ICMLA). 2020;1317–1323.
31. Al Helali B, Chen Q, Xue B, et al. *Multi-tree genetic programming-based transformation for transfer learning in symbolic regression with highly incomplete data*. 2020 IEEE Congress on Evolutionary Computation (CEC). 2020;1–8.
32. Al Helali B, Chen Q, Xue B, et al. *Multi-tree genetic programming for feature construction-based domain adaptation in symbolic regression with incomplete data*. Proceedings of the 2020 Genetic and Evolutionary Computation Conference. 2020;913–921.
33. Al Helali B, Chen Q, Xue B, et al. Multitree genetic programming with new operators for transfer learning in symbolic regression with incomplete data. *IEEE Transactions on Evolutionary Computation*. 2021;25(6):1049–1063.
34. Sanodiya RK, Mathew J, Saha S, et al. Particle swarm optimization based parameter selection technique for unsupervised discriminant analysis in transfer learning framework. *Applied Intelligence*. 2020;50:3071–3089.
35. Chen Q, Xue B, Zhang M. *Instance based transfer learning for genetic programming for symbolic regression*. 2019 IEEE Congress on Evolutionary Computation (CEC). 2019;3006–3013.
36. Friedjungová M, Jirina M. Asymmetric Heterogeneous Transfer Learning: A Survey. *DATA*. 2017;17–27.
37. Peng X, Huang Z, Sun X, et al. *Domain agnostic learning with disentangled representations*. International Conference on Machine Learning. 2019;97:5102–5112.
38. Nyathi T, Pillay N. *On the transfer learning of genetic programming classification algorithms*. In Theory and Practice of Natural Computing: 10th International Conference. Tsukuba. Japan. 2021;15:47–58.
39. Gong B, Shi Y, Sha F, et al. Geodesic flow kernel for unsupervised domain adaptation. *2012 IEEE conference on computer vision and pattern recognition*. 2012;2066–2073.
40. Sim T, Baker S, Bsat M. The CMU pose, illumination, and expression (PIE) database. In *Proceedings of fifth IEEE international conference on automatic face gesture recognition*. 2012;53–58.
41. Vergara A, Vembu S, Ayhan, T, et al. Chemical gas sensor drift compensation using classifier ensembles. *Sensors and Actuators B: Chemical*. 2012;166:320–329.
42. Hull JJ. A database for handwritten text recognition research. *IEEE Transactions on pattern analysis and machine intelligence*. 1994;16(5):550–554.
43. Deng L. The mnist database of handwritten digit images for machine learning research [best of the web]. *IEEE signal processing magazine*. 2012;29(6):141–142.
44. Ben Dor A, Shamir R, Yakhini Z. Clustering gene expression patterns. *J Comput Biol*. 1999;6(3):281–297.
45. Pérez Núñez P, Luaces O, Díez J, et al. *Tripadvisor restaurant reviews*. Zenodo. 2021.