Evolutionary Transfer Optimization—A New Frontier in Evolutionary Computation Research

Kay Chen Tan
The Hong Kong Polytechnic University, HONG KONG SAR

Liang Feng
Chongqing University, CHINA

Min Jiang
Xiamen University, CHINA

Abstract—The evolutionary algorithm (EA) is a nature-inspired population-based search method that works on Darwinian principles of natural selection. Due to its strong search capability and simplicity of implementation, EA has been successfully applied to solve many complex optimization problems, which cannot be easily solved by traditional exact mathematical approaches, such as linear programming, quadratic programming, and convex optimization. Despite its great success, it is worth noting that traditional EA solvers start the search from scratch by assuming zero prior knowledge about the task at hand. However, as problems seldom exist in isolation, solving one problem may yield useful information for solving other related problems. There has been growing interest in conducting research on evolutionary transfer optimization (ETO) in recent years: a paradigm that integrates EA solvers with knowledge learning and transfer across related domains to achieve better optimization efficiency and performance. This paper provides an overview of existing works of ETO based on the type of problems being solved by these methods, which are *ETO for Optimization in Uncertain Environment*, *ETO for Multitask Optimization*, *ETO for Complex Optimization*, *ETO for Multi/Many-Objective Optimization*, and *ETO for Machine Learning Applications*. The paper also highlights some of the challenges faced in this emerging research field of computational intelligence and discusses some promising future research directions in ETO. It is hoped that the study presented in this paper can help to inspire the development of more advanced ETO methods and applications.

Digital Object Identifier 10.1109/MCI.2020.3039066
Date of current version: 12 January 2021

I. Introduction

Evolutionary algorithms (EAs) are population-based search approaches that are inspired by the principles of natural selection and genetics [1]. To solve a given optimization problem, an EA typically starts with a randomly generated population of individuals using a problem-specific encoding scheme. The individuals then undergo genetic operations, i.e., crossover, mutation, and selection, to produce a

Corresponding Author: Liang Feng (liangf@cqu.edu.cn).

 1556-603X/21©2021IEEE

©SHUTTERSTOCK.COM/ELLAGRIN

new generation of offspring iteratively, until a predefined condition is satisfied. In contrast to exact mathematical optimization methods, e.g., linear programming [2], quadratic programming [3], and least squares [4], EA has strong search capability in solving complex optimization problems that are non-convex or multimodal in nature. Over the years, EA has attracted much attention in both academia and industry, and has been successfully applied to many applications, including problems with single objective optimization [5], multi-objective optimization [6], and dynamic optimization [7].

Despite the success enjoyed by EAs, it is worth noting that existing EA solvers usually conduct the search process from scratch, independent of how similar the new problem encountered is to those already solved in the past [8]. Therefore, existing EAs often do not learn from problems, and the search capabilities of EA solvers do not grow with problem-solving experiences automatically. However, in practice, problems seldom exist in isolation. Ignoring the search experience of previous optimization processes on related problems may lead to unnecessary computational costs in repeated searches on similar problems. The ability to generalize what has been learned across problems is likely to have a great impact in solving the ever-increasing complexity of real-world optimization problems. Keeping these in mind, the development of new learning and transfer approaches for creating a new class of "learnable" EAs is thus desirable.

In machine learning, the idea of taking advantage of useful traits across problem domains towards enhanced learning performance has received significant interest under the term of transfer learning (TL) in recent years [9], [10]. TL leverages the knowledge gained from source domains, where much high-quality training data are available, to improve the learning model in a related target domain containing fewer high-quality training data [9]. However, existing research studies on TL have largely been confined to machine learning applications, such as computer vision [11], natural language processing [12], and speech recognition [13], where problem or application data are usually widely available for knowledge learning and transfer. In the context of evolutionary optimization, usually only a given objective function and limited problem-specific data are available a-priori, and new and advanced knowledge learning and transfer approaches are thus needed for the development of "learnable" EAs.

There is increasing interest in conducting research on evolutionary transfer optimization (ETO), which is a paradigm that integrates EA solvers with knowledge learning and transfer across related domains to achieve better optimization efficiency and performance. It is noted that the notion of memetics is closely related to that of knowledge representations for

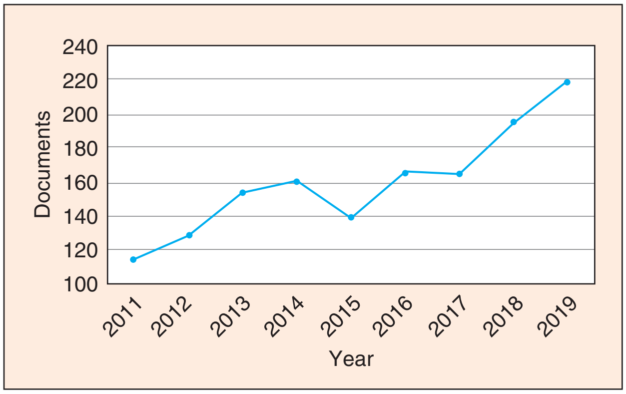

FIGURE 1 Number of published research papers (recorded in Scopus) related to ETO.

problem-solving [14]. As depicted in Fig. 1, by a simple search of the terms "Evolutionary Transfer Optimization" and "Evolutionary Transfer Algorithm" in Scopus[1], the number of published research papers increases from approximately 110 in year 2011 to approximately 220 in year 2019. Note that there may be publications within the scope of ETO that are missing in the statistics, due to the simple criteria used in the search. The ETO aims to improve the performance of EA solvers in terms of the quality of solution and speed of search by learning and transferring useful traits across related problems in the form of solutions, structured knowledge, etc. In the literature, various ETO approaches have been proposed to solve both benchmark and practical optimization problems, including dynamic optimization problems [15], multi-objective optimization problems [16], multitask optimization problems [17], combinatorial optimization problems [18], and expensive optimization problems [19]. However, to the best of our knowledge, there is no or little effort made to discuss and categorize existing ETO approaches. Moreover, it is observed that existing ETO research progresses disjointedly, which may impede the development of new methodologies and applications in this emerging field. This paper thus presents an exposition of existing ETO approaches according to the type of problems being solved. It also highlights some challenges faced by the ETO and discusses a number of future research directions. We hope this paper will attract attention from researchers in both evolutionary computation and machine learning communities to further design advanced ETO approaches to address the ever-increasing complexity of optimization problems.

The remainder of this paper is organized as follows. Section II presents the background of ETO, which includes the introduction of evolutionary optimization and transfer learning. Section III gives a review of existing state-of-the-art ETO approaches. Discussions on future research directions of ETO are given in Section IV. Conclusions are drawn in Section V.

II. Background

This section provides a brief review of EAs, including the definition and workflow of a generic EA. It also introduces transfer learning and discusses how TL can be integrated with EAs to achieve better optimization performance.

A. Evolutionary Algorithm

EAs denote a class of population-based metaheuristic optimization algorithms that are inspired by biological evolution, such as reproduction, mutation, and selection [1]. It has been applied to solve many problems that cannot be easily addressed by traditional approaches, such as NP-hard combinatorial optimization problems [20], optimization problems with a large number of decision variables [21], and optimization problems with conflicting objectives [22]. Typically, upon deciding on an encoding representation, e.g., binary encoding [23], gray encoding [24], and real number encoding [25], the algorithm starts with a randomly generated initial population. After fitness evaluation of all individuals in the population, selection is performed to identify fitter individuals to undergo reproduction for generating the offspring. Then, genetic operations such as crossover and mutation will be carried out [25]. Such an evolutionary process will be executed iteratively until certain predefined stopping criteria are satisfied.

From the workflow of EAs, we can observe that the optimization process is based on the movement of individuals in a population within the search space of the problem of interest. Therefore, the population in an evolutionary search contains key information for solving the problem, and useful traits could be learned from moving traces of the population and be transferred across problem domains if applicable to guide the search towards enhanced optimization performance.

B. Transfer Learning

It is known that humans naturally transfer knowledge between tasks, such as by recognizing and applying relevant knowledge from previous problem-solving experiences when a new task is encountered [26]. Usually the more relevant a new task is to those already solved in the past, the easier the new task can be solved efficiently. In the context of computer science, transfer learning has been used to leverage the knowledge learned in one or more source tasks to improve the learning performance of a related target task. As depicted in Fig. 2, the tasks can be in the area of traditional machine learning, such as classification, regression, reinforcement learning, and deep learning [27], [28]. The knowledge learned and transferred across tasks can come from diverse representations [29], e.g., neural knowledge, matrix knowledge, and tree knowledge, which is often problem dependent. TL has attracted increasing attention in recent years, and many algorithms have been proposed to improve the learning performance of TL in various applications, such as face recognition, indoor localization, and deep learning training [30].

The success of TL in machine learning shows that it is capable of reducing the effort needed to build a model from scratch by using fundamental logic or structured knowledge within one domain and applying it to another. To achieve better optimization performance, instead of starting an evolutionary search from scratch, TL can be integrated into EAs to generalize what has been optimized in the past. However, the

[1] https://www.scopus.com/

design of effective TL in EAs requires new representations of knowledge as well as new learning methods for knowledge transfer across tasks, since usually only objective function and limited problem-specific data are available in a-prior to an evolutionary optimization process.

III. Evolutionary Transfer Optimization

The study of ETO has attracted increasing interest in the field of computational intelligence. As shown in Fig. 3, existing ETO approaches can be categorized into *homogenous ETO* and *heterogeneous ETO* from the perspective of algorithmic design. The former focuses on knowledge transfer in evolutionary search across problems sharing a common search space, while the latter considers learning and transfer of knowledge across problems possessing diverse search spaces, such as having different numbers of dimensions, different decision variables, and different objectives.

As shown in Fig. 3, based on the type of problems being solved, existing ETO studies can be categorized as follows: 1) ETO for Optimization in Uncertain Environment; 2) ETO for

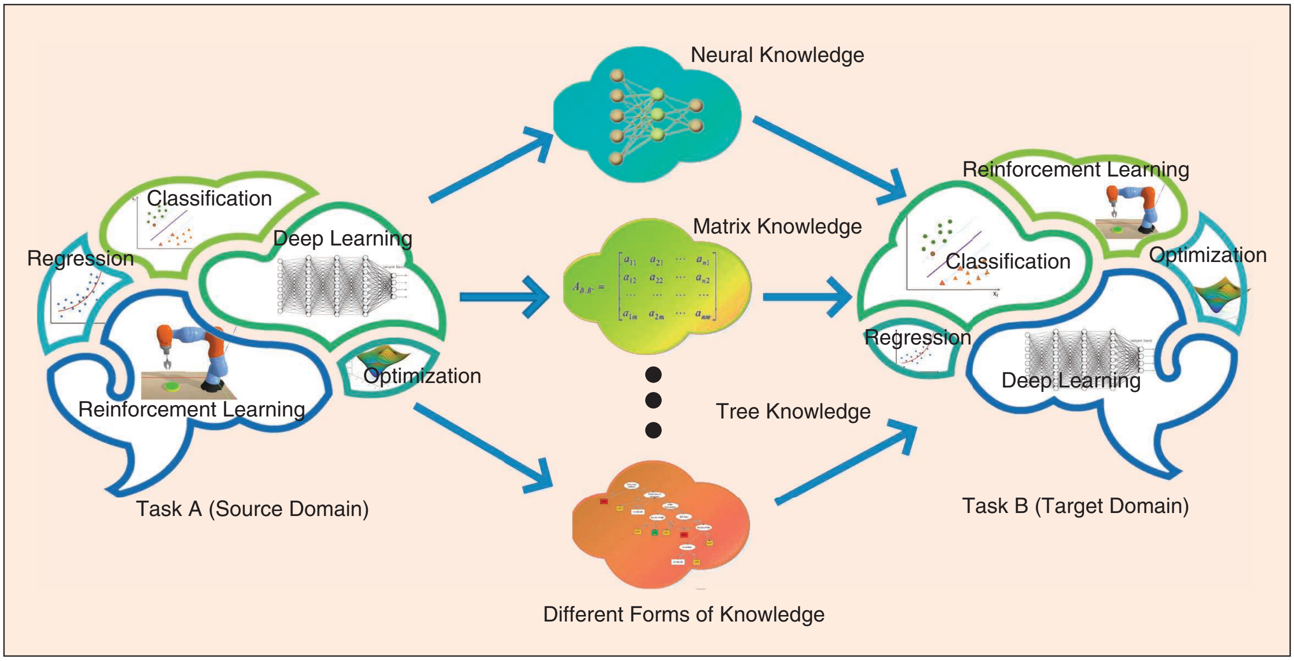

FIGURE 2 Illustration of transfer learning.

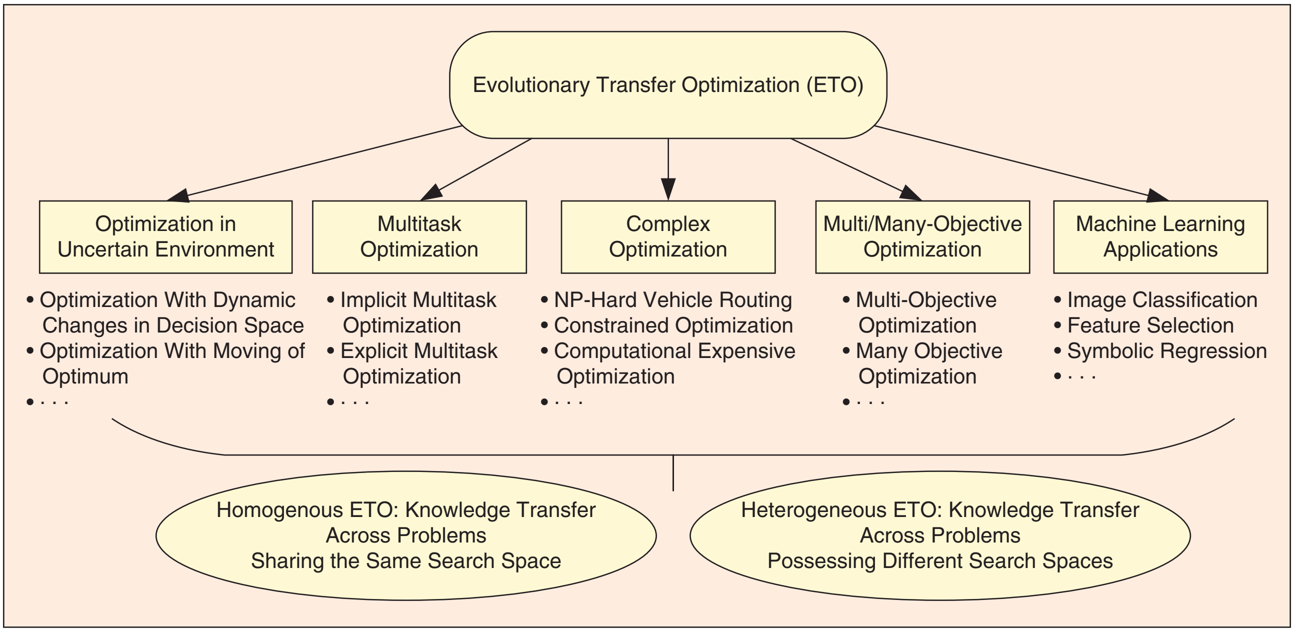

FIGURE 3 Category of existing ETO approaches.

Multitask Optimization; 3) ETO for Complex Optimization; 4) ETO for Multi/Many-Objective Optimization; 5) ETO for Machine Learning Applications.

A. ETO for Optimization in Uncertain Environment

In practical optimization problems, evaluation of a solution is often subject to different uncertainties, such as noise or approximations in function evaluation, dynamic change of decision variables and/or fitness functions, and robustness for the solutions [31], [32]. In an uncertain environment, the optimal solution found at a particular time instance may not necessarily be the optimal solution at another time instance. As illustrated in Fig. 4, the global optimum for $f(t)$ at time instance t_1 is no longer the global optimum at time instance t_2. In contrast, the local optimum achieved at t_1 is actually closer to the global optimum at t_2. Although the problem changes dynamically, the dynamic nature at two adjacent time instances may share certain similarities. Therefore, the search experience obtained before the dynamic change can provide useful knowledge for enhancing the problem-solving for the newly changed problem when properly harnessed.

A number of studies on ETO for optimization in uncertain environments have been conducted recently. In [15], Jiang *et al.* improved the prediction of moving optima in dynamic optimization by considering the nonindependent and identically distributed nature of data obtained during the evolutionary search process via transfer component analysis. In [33], to solve a dynamic vehicle routing problem, Zhou *et al.* proposed an approach to capture the structured knowledge from optimized routing solution in the early time slot, which is then reused to bias customer-vehicle assignment in the evolutionary search when a change is detected. In [34], Hatzakis and Wallace proposed a time-series model with an EA to predict the location of the next optimal solution after a change occurs. In their approach, a sequence of optimal positions found in the past are used as knowledge in the construction of a prediction model. Similar ideas have been proposed in [35], [36], and [37], where movements of optima are predicted with knowledge learned from past search experiences using kernel-based methods, manifold transfer learning, knee point-based imbalanced learning, etc.

As data in uncertain environments may be noisy and/or dynamic, specific ETO designs are often needed to identify and handle such data before any knowledge transfer. It also remains a challenge to develop robust and incremental transfer learning methods for positive knowledge transfer [9], while the evolutionary search progresses online.

B. ETO for Multitask Optimization

The multitask optimization problem (MTOP) focuses on solving multiple self-contained tasks simultaneously. As depicted in Fig. 5, the input of multitask optimization is a set of optimization tasks or problems, and the output is a set of optimal solutions, each corresponding to a task. In MTOP, each task is an independent optimization problem that possesses a unique problem search space, which can be either a single- or multi-objective optimization problem [17]. The MTOP is inspired by humans' capability of performing multiple tasks with apparent simultaneity and the well-established concept of multitask learning in predictive analytics. In particular, by transferring useful knowledge across tasks online as the evolutionary optimization progresses, the solving of one problem may lead to the related problem being solved automatically, which may thus enhance the search capability in terms of both efficiency and performance.

Various ETO methods have been proposed for solving MTOPs. In [17], Gupta *et al.* presented a multifactorial evolutionary algorithm (MFEA) for solving MTOPs with a unified representation. The knowledge transfer across tasks in this work is realized by crossover. In [38], Ding *et al.* improved MFEA by proposing decision variable translation and decision variable shuffling to facilitate knowledge transfer between optimization problems having different locations of optima and different numbers of decision variables, respectively. Towards positive knowledge transfer, Bali *et al.* [39] improved the MFEA by proposing a linearized domain adaptation strategy to transform the search space of a simple task to the search space similar to its constitutive complex task. Other improvements or variants of MFEA include [16], [40]–[46]. In addition to using genetic crossover for knowledge transfer, [47] derived a single-layer autoencoder to transfer the optimized solutions in one task to another via matrix multiplication. A similar idea

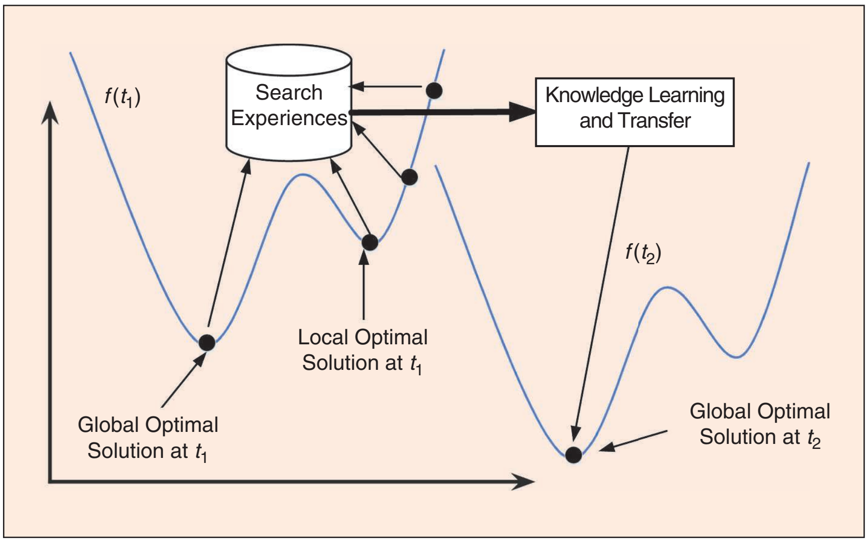

FIGURE 4 Example of ETO for problem-solving in an uncertain environment, e.g., dynamic change of global optimum.

was proposed in [18], where a weighted l_1 norm-regularized learning process is formulated to adapt customer distributions across vehicle routing problems for the transfer of high-quality routing solutions during the search process. Through an independent transfer component, the design of learning and transfer of knowledge can be developed in ETO for enhanced multitask optimization performance. In [48], Ma *et al.* introduced a two-level transfer learning method for evolutionary multitasking, for which the upper-level implements inter-task knowledge transfer learning via genetic crossover and the lower-level performs intra-task transfer learning based on information transfer of decision variables.

It is worth noting that the correlation between tasks is essential to achieve positive knowledge transfer for MTOPs. As discussed in Section II-A, the search in EAs is an iterative process, and the guidance of the search towards areas of high-quality solutions is time-dependent. Therefore, the usefulness of knowledge transfer across tasks may vary at different stages of the evolutionary search, i.e., the evaluation of correlation in MTOPs is dynamic rather than a static problem. To utilize the availability of today's cloud computing, the design of ETO approaches capable of solving a large number of tasks simultaneously is also an interesting topic to be explored.

C. ETO for Complex Optimization

Many real-world applications involve complex optimization problems, e.g., non-convex problems (e.g., Fig. 6(a)), problems possess many constraints (e.g., Fig. 6(b) and Fig. 6(d)), problems cannot be solved in polynomial time (e.g., Fig. 6(b)), problems contain many local optima (e.g., Fig. 6(c)), and extremely computationally expensive problems (e.g., Fig. 6(d)). To address the complexity of these problems, many features have been proposed in EAs to improve the search efficiency, such as new search operators [20], adaptive mechanisms [49], and new search space construction [50].

By learning and transferring useful knowledge from related and simpler problem domains, ETO can help to deal with complex optimization problems. In [51], Louis *et al.* presented a study to acquire problem specific knowledge to aid the search in a genetic algorithm (GA) via case-based reasoning. Instead of starting anew on each problem, appropriate intermediate solutions drawn from similar problems solved previously are periodically injected into the GA population. In [52], Cunningham and Smyth applied knowledge transfer in the form of established high-quality schedules from past problems to bias the search for new traveling salesman problems (TSPs). In addition to reusing past optimized solutions, Roberto *et al.* [53] proposed the transfer of structural information from subproblems to bias construction of the aggregation matrix of the estimation of distribution algorithm (EDA) for solving a multi-marker tagging single-nucleotide polymorphism (SNP) selection problem. Based on semidefinite programming, [54] and [55] proposed a homogeneous and heterogeneous transfer learning approach to speed up the evolutionary optimization of vehicle routing via knowledge transferred from the solved vehicle routing and arc routing problems, respectively. Chaabani *et al.* [56] integrated a co-evolutionary decomposition-based algorithm with transfer learning to enhance the EA search for solving bi-level production-distribution problems in supply chain management. Tan *et al.* [19] proposed an adaptive knowledge transfer framework for surrogate-assisted evolutionary search of computationally expensive problems based on the idea of multiproblem surrogates, which enables EAs to acquire and spontaneously transfer the learned models across problems towards efficient global optimization.

From these approaches, we can observe that the success of ETO for complex optimization relies greatly on finding simpler and related problems of the encountered complex

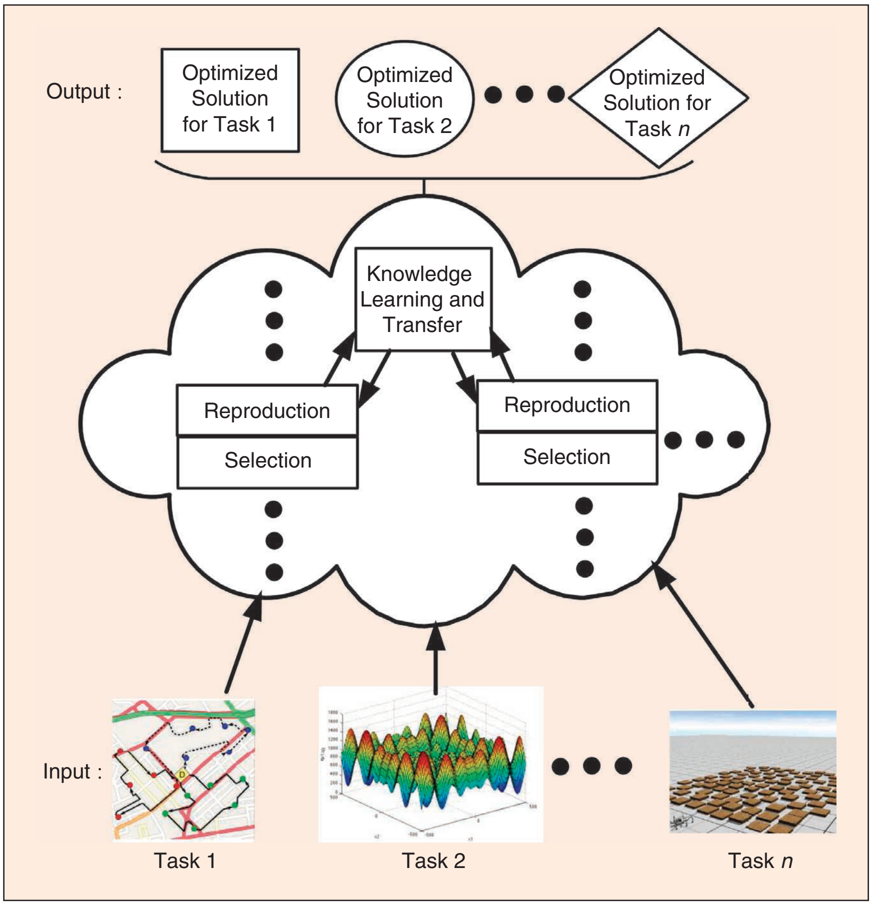

FIGURE 5 Illustration of multitask optimization.

problem. This is challenging since there is no universal database containing different types of optimization problems and their characteristics. Instead of finding a similar problem, an alternative approach is to independently and artificially construct useful and simpler problems in ETO based on characteristics of the given complex problem.

D. ETO for Multi/Many-Objective Optimization

The multi-objective problem (MOP) contains several conflicting objectives to be optimized simultaneously [22], [57]. Mathematically, an MOP can be stated as follows:

$$\begin{aligned} \min/\max \quad & F(\mathbf{x}) = \left(f_1(\mathbf{x}), f_2(\mathbf{x}), \ldots, f_d(\mathbf{x}) \right) \\ s.t. \quad & g_j(\mathbf{x}) \geq 0, j = 1, 2, \cdots, J \\ & h_k(\mathbf{x}) = 0, k = 1, 2, \cdots, K \\ & x_i^L \leq x_i \leq x_i^U, i = 1, 2, \cdots, n \end{aligned} \tag{1}$$

where $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$ is the decision vector, $g_j(\cdot)$ denotes the inequality constraint, and $h_k(\cdot)$ gives the equality constraint. d is the number of objective functions, J is the number of inequality constraints, and K is the number of equality constraints. x_i^L and x_i^U are the lower and upper bound of x_i, respectively. Many-objective problems usually refer to MOPs having more than three

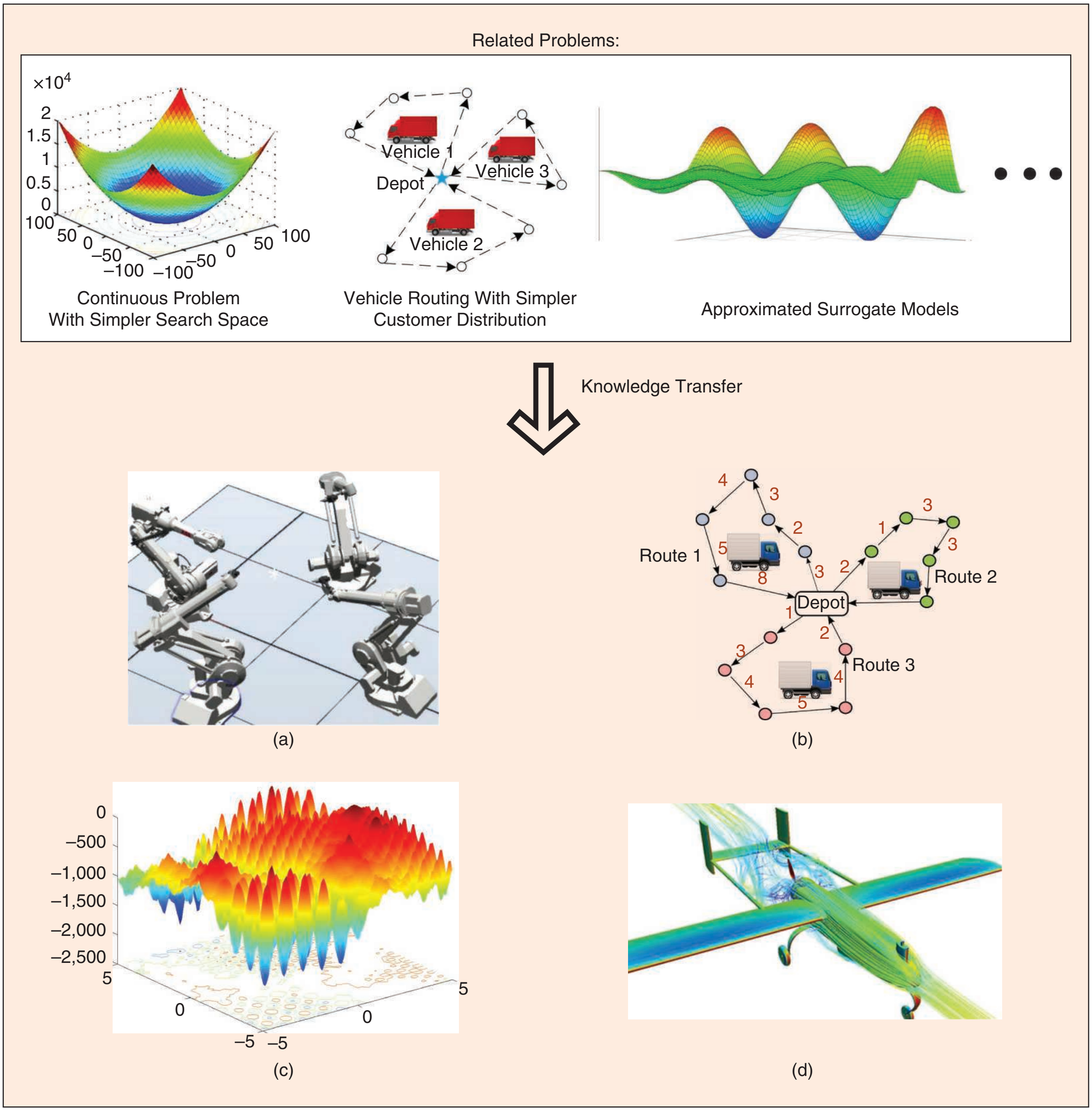

FIGURE 6 Examples of complex optimization problems. (a) Non-convex optimization problem, (b) NP-hard optimization problem, (c) Optimization problem with may local optima, and (d) Computational expensive optimization problem.

objectives (i.e., $d > 3$) [6]. In Eq. 1, the final solution of an MOP is not a single point but rather a set of Pareto-optimal solutions. Compared to single-objective optimization, solving MOPs is more challenging due to the increase in computational time and search space when the number of objectives is increased.

ETO has been applied to multi- and many-objective optimization problems via knowledge transfer across problems. As depicted in Fig. 7, useful knowledge in multi- and many-objective optimization can be in the form of non-dominated solutions, surrogated models, etc. Yang *et al.* [58] proposed an ETO algorithm for solving MOPs, which is assisted by two surrogates: one aims to guide the algorithm to quickly find a promising subregion, while the other leverages good solutions according to knowledge transferred from the first surrogate. In [59], Feng *et al.* presented a single-layer denoising autoencoder to transfer knowledge in the form of MOP solutions from past MOPs when a new MOP is encountered. Lin *et al.* [60] considered the concept of dominance in knowledge transfer by proposing an effective solution transfer approach for solving MOPs via evolutionary multitasking. Liang *et al.* [61] proposed a two-stage adaptive knowledge transfer method by considering population distribution for both multi- and many-objective optimization using evolutionary multitasking.

Intuitively, the number of non-dominated solutions to be transferred will grow dramatically as the number of objectives is increased in many-objective optimization problems. The construction of proper mapping and the learning and transfer of useful knowledge between source and target domains are thus more challenging compared to ETO for single objective optimization. Deeper analysis of what is regarded as useful information and how such information is related to the multiple objectives across problem domains can help to better understand how and when ETO performs well in multi- and many-objective optimization.

E. ETO for Machine Learning Applications

Both optimization and learning play important roles in machine learning applications. As illustrated in Fig. 8(a), the loss function can be formulated as a minimization problem based on the classification error [62]–[64]. As shown in Fig. 8(b), the minimization of a loss function based on the prediction error and regularization value for the training data can be used to represent the objective function of regression or time-series prediction applications [65]. As illustrated in Fig. 8(c), another example is feature selection, which aims to extract a subset of features by maximizing the learning performance of the machine learning task of interest [66].

A number of ETO approaches have been proposed for machine learning applications by leveraging useful knowledge across learning problem domains. Iqbal *et al.* applied transfer learning in genetic programming to learn and transfer useful traits from simple to complex problem domains for multi-class texture classification and image classification in [67] and [68], respectively. In [69], Chen *et al.* considered the generalization and domain adaptation of genetic programming in the context of symbolic regression. In [70], Iqbal *et al.* identified the building blocks of knowledge in a learning classifier system to improve the performance of EAs across Boolean problems. Wei *et al.* [71] extended gene expression programming for multi-class classification by decomposing a problem into several binary classification subtasks. Evolutionary transfer optimization was then performed on all the subtasks concurrently. In [66], it is shown that the transfer of domain knowledge in evolutionary computation can lead to more efficient feature selection for real-world applications, such as in financial distress prediction, handwritten digit recognition, and biomarker detection. More recently, Mouret *et al.* [72] conducted online knowledge transfer across optimization problems of different robot movements. Martinez *et al.* [73] proposed an approach of simultaneously evolving several deep Q-learning models for interrelated reinforcement learning tasks.

Although some encouraging results have been reported, existing ETO approaches mainly focus on solving small- to medium-scale problems in machine learning applications. With the availability of big data and rapid development of modern hardware technologies, it is useful to develop advanced ETO algorithms capable of leveraging these resources and hardware to address today's ever-growing range and scale of demands in machine learning applications, e.g., fast and automatic deep neural network configurations for image processing, speech recognition, etc.

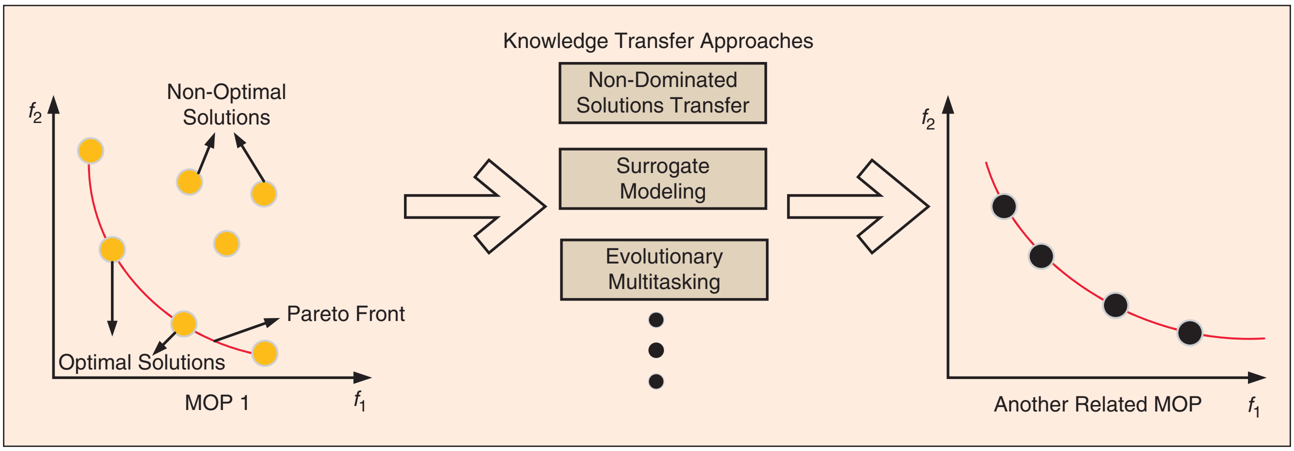

FIGURE 7 Illustration of knowledge transfer across MOPs.

IV. Future Research Directions

Although ETO has made remarkable progress in recent years, there are a number of potential research directions to be explored, as discussed below.

A. ETO for Large-Scale Optimization

With the rapid development of digital technologies and advanced communications, large-scale optimization problems are encountered in many real-world applications, such as package delivery for e-commerce, fraud detection, and social media analysis [74]. It is a challenge to solve large-scale optimization problems, since a large number of decision variables will lead to a very large search space, resulting in a large computational cost [75]. To improve the scalability of EAs for large-scale optimization, different approaches have been proposed to simplify the search space or to improve the search capability of EAs, such as dimension reduction [76], cooperative coevolution [77], and new search mechanisms [78].

In contrast to existing EAs, ETO can bring a new perspective for enhancing the scalability of EAs in solving large-scale problems. In particular, as mentioned above, problems seldom exist in isolation. With a proper design of ETO approaches, the success of EAs in solving small- and/or medium-size problems can be leveraged for solving related large-scale problems. Possible research topics include 1) the design of ETO algorithms for solving large-scale optimization problems by learning from related problems; 2) the design of ETO algorithms capable of conducting evolutionary searches for large-scale problems in a reduced knowledge space; and 3) the design of ETO algorithms for conducting searches in the large-scale problem space and the reduced problem space concurrently so that useful traits found in the simplified problem space can be transferred to guide the evolutionary search for the large-scale problem.

B. ETO for Multi-Form Optimization

In contrast to existing ETO approaches that deal with distinct optimization problems, multi-form optimization exploits alternate formulation(s) of a single target optimization problem [79]. As different formulations of a single optimization problem may capture different properties or landscapes of the problem

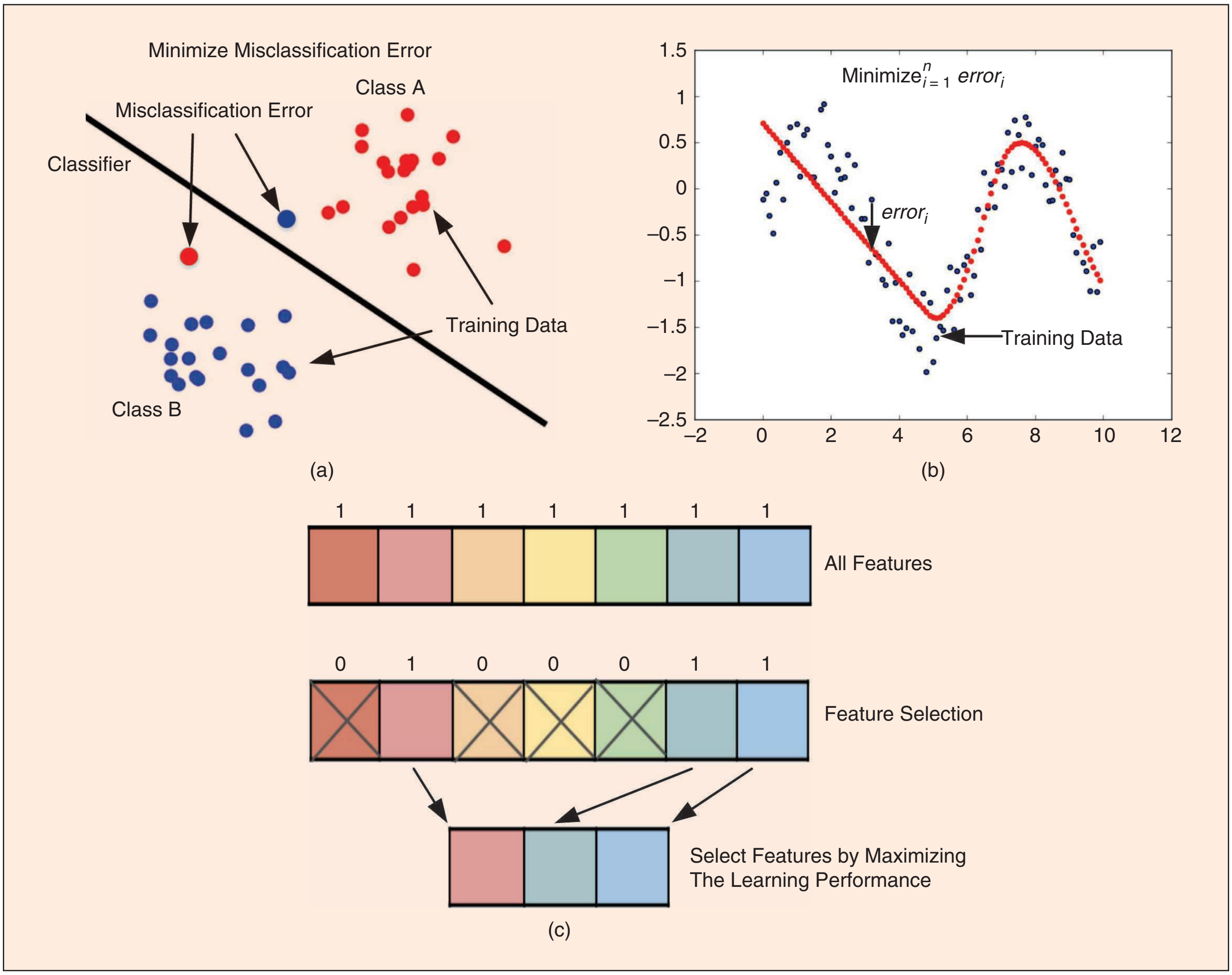

FIGURE 8 Illustration of classification, regression, and feature selection problem. (a) Classification problem, (b) regression problem, and (c) feature selection.

encountered, search experiences obtained from these alternate formulations may lead to better problem-solving for the target problem [26]. Instead of performing evolutionary search on the original problem formulation with a complicated search space, the search may be conducted based on alternate formulations containing simpler search spaces. By transferring useful knowledge obtained in the simpler search space back to the original space, the search can be guided to regions of high-quality solutions for better optimization performance.

In practice, there are several ways to construct different formulations of a given optimization problem, such as changing the constraints, altering the structure of objective function(s), and reducing the number of control variables. However, it is often difficult to ascertain whether the constructed formulation is useful. Moreover, it is also a challenge to conduct knowledge transfer across different formulations effectively. Potential research topics include the following:

- Design of useful alternate problem formulations in the context of single- and multi-objective optimization, constrained optimization, etc.
- Design of ETO algorithms capable of automatically generating and configuring different formulations of a given problem.
- Efficient allocation of computational resources for conducting evolutionary search based on different problem formulations.

C. ETO for Deep Learning

Deep learning has been applied to many real-world applications, such as image classification [63], speech recognition [80], and recommendation systems [81]. However, the success of deep learning often relies on the "deep" architecture of neural networks to learn representations from data such as images, video or text automatically, without introducing hand-coded rules or human domain knowledge [63]. Currently, state-of-the-arts deep neural networks are often manually designed with expertise from deep learning and problem-to-be-solved.

To make deep learning more accessible and to reduce the need for human expertise, a number of dedicated EAs have been proposed for the automatic design of deep neural architectures. In [82], Zhang *et al.* integrated multi-objective evolutionary search with ensemble learning to optimize the configuration of a deep belief network. In [83], Liu *et al.* introduced a layerwise method based on multi-objective evolutionary optimization for structural learning of deep neural networks. Sun *et al.* applied a genetic algorithm to automatically design a deep convolutional neural network for image classification [64]. Traditional EAs usually require a large amount of computational cost to find an acceptable deep neural architecture. On the other hand, the transfer of knowledge, such as neural architectures [84], deep features [85], and internal representations [86], across deep learning models can often improve the deep learning performance. Designing effective ETO methods for automatic deep learning is thus a potential research direction. Such ETO algorithms should learn and transfer success from well-designed deep learning models to enhance the evolutionary search of new deep learning models.

Potential research topics include 1) the design of ETO algorithms for deep neural network architecture searches with knowledge transfer across neural network models in the form of small neural nets, learned deep features, etc.; 2) the design of ETO algorithms for fast hyper-parameters optimization in deep neural network architectures based on hyper-parameters of trained networks across problem domains; 3) the design of ETO approaches with unified knowledge space across different deep learning models; and 4) the design of graphics processing unit (GPU)-based ETO approaches for fast and efficient automatic design of deep neural network architectures.

D. ETO in Complex Data Environment

Modern applications often contain data with complex properties, such as noise, imbalanced information, and limited labels. For instance, the tag noise in an image may be generated by unreliable labeling using a fast framework such as Amazon Mechanical Turk [87]. The breast cancer dataset usually contains a large number of benign samples and a small number of malignant samples, leading to an imbalance classification problem [88]. In logistics, there may exist vehicle routing tasks involving customers ranging from hundreds to thousands, and often there is no label information to relate one customer to another [89]. As improper knowledge transfer across problems is undesirable, designing ETO approaches to handle complex data properties for positive knowledge transfer is important. A number of approaches have been proposed for knowledge transfer in realistic scenarios with complex data properties. In particular, Ge *et al.* considered the issue of negative transfer and imbalanced data distributions in [90], while Yu *et al.* focused on TL designs in cases where label noises exist [91]. Cao *et al.* presented a study on adaptive TL design, which can automatically estimate the similarity between source and target tasks [92].

Potential research topics for ETO in complex data environment include 1) the design of robust ETO algorithms for positive knowledge transfer from noisy data; 2) the design of online ETO algorithms for positive knowledge transfer across problem domains where the data for learning and transfer appear in sequential or batch order; 3) the design of ETO algorithms for positive knowledge learning and transfer across problems having imbalanced data or without labels; and 4) the design of ETO approaches for efficient knowledge transfer in cases where the problem property changes in dynamic environments.

E. Theoretical Study of ETO

Despite the increasing interest in designing new ETO algorithms, there is a lack of rigorous theoretical studies on ETO. Although there are some theoretical analyses of knowledge transfer between classification tasks [93], [94], such analyses do not directly apply to EAs. There are a few studies on the similarity measure between optimization problems for positive knowledge transfer in evolutionary search [95]–[98]. To have a better understanding of how and when ETO works, more theoretical studies and analyses of ETO are necessary. Possible topics of research include the following:

❏ Theoretical study on whether knowledge acquired in a particular source problem can improve the evolutionary search in a target problem (for a specific type of optimization problem).
❏ Theoretical analysis of the correlation between the amount of knowledge transferred and the improvement achieved in the target problem.
❏ Definition on optimal inter-task mapping and how transfer efficacy can be impacted by the constructed inter-task mapping.
❏ Definition on useful representation of knowledge that can be transferred across heterogeneous problem domains.

V. Conclusions

ETO is a generic and emerging paradigm that integrates EA solvers with knowledge learning and transfer across related domains to achieve efficient and better optimization performance. In this paper, we have presented an overview of existing ETO approaches based on the type of problems being solved, which includes ETO for Optimization in Uncertain Environment, ETO for Multitask Optimization, ETO for Complex Optimization, ETO for Multi/Many-Objective Optimization, and ETO for Machine Learning Applications. This study shows that with knowledge learned and transferred across related problem domains, the search and optimization capability of EAs can be improved, when properly harnessed.

As one of the emerging research areas in computational intelligence, there are many challenges and open research questions in ETO. We have discussed some of the challenges and potential research directions, such as ETO for large-scale optimization, ETO for multi-form optimization, ETO for deep learning, ETO in complex data environment, and theoretical study of ETO. Since problems often share certain correlations in nature, it is believed that ETO can become an important technology for tackling the ever-increasing complexity and interdependency of optimization problems currently encountered in real-world applications.

Acknowledgment

This work is partially supported by the National Key Research and Development Project, Ministry of Science and Technology, China (Grant No. 2018AAA0101301), the National Natural Science Foundation of China (NSFC) under Grant No. 61876162, 61876025, and 61673328, the Venture & Innovation Support Program for Chongqing Overseas Returnees under Grant No. cx2018044 and cx2019020, and the Research Grants Council of the Hong Kong SAR under grant No. CityU11202418 and No. CityU11209219.

References

[1] R. Salomon, "Evolutionary algorithms and gradient search: similarities and differences," *IEEE Trans. Evol. Comput.*, 2(2):45–55, 1998. doi: 10.1109/4235.728207.
[2] G. B. Dantzig and N. N. Thapa, *Linear Programming 1: Introduction*. Berlin: Springer-Verlag, 1997.
[3] Z. Dostl, *Optimal Quadratic Programming Algorithms: With Applications to Variational Inequalities*, 1st ed. Springer-Verlag, 2009.
[4] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge Univ. Press, 2004.
[5] M. M. Islam, H. K. Singh, and T. Ray, "A surrogate assisted approach for single-objective bilevel optimization," *IEEE Trans. Evol. Comput.*, 21(5):681–696, 2017. doi: 10.1109/TEVC.2017.2670659.
[6] B. Li, J. Li, K. Tang, and X. Yao, "Many-objective evolutionary algorithms: A survey," *ACM Comput. Surv.*, 48(1):1–35, 2015. doi: 10.1145/2792984.
[7] S. Yang and X. Yao, Eds. *Evolutionary Computation for Dynamic Optimization Problems* (Studies in Computational Intelligence). Berlin: Springer-Verlag, 2013.
[8] Y. Jin, Ed. *Knowledge Incorporation in Evolutionary Computation*. Berlin: Springer-Verlag, 2005.
[9] S. J. Pan and Q. Yang, A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.*, 22(10):1345–1359, 2010. doi: 10.1109/TKDE.2009.191.
[10] A. Ng. 2016. https://nips.cc/conferences/2016/tutorialsintros
[11] C. Sferrazza and R. DAndrea, "Transfer learning for vision-based tactile sensing," in *Proc. IEEE/RSJ Int. Conf. Intell. Robots Syst. (IROS)*, 2019, pp. 7961–7967. doi: 10.1109/IROS40897.2019.8967571.
[12] S. Ruder, M. E. Peters, S. Swayamdipta, and T. Wolf, "Transfer learning in natural language processing," in *Proc. 2019 Conf. North Amer. Chap. Assoc. Comput. Linguist: Tuts.*, pp. 15–18. doi: 10.18653/v1/N19-5004.
[13] J. Kunze, L. Kirsch, I. Kurenkov, A. Krug, J. Johannsmeier, and S. Stober, "Transfer learning for speech recognition on a budget," in *Proc. 2nd Workshop Represent. Learn. NLP*, 2017, pp. 168–177. doi: 10.18653/v1/W17-2620.
[14] A. Gupta and Y. S. Ong, *Memetic Computation: The Mainspring of Knowledge Transfer in a Data-Driven Optimization Era*. Springer-Verlag, 2019.
[15] M. Jiang, Z. Huang, L. Qiu, W. Huang, and G. G. Yen, "Transfer learning-based dynamic multiobjective optimization algorithms," *IEEE Trans. Evol. Comput.*, 22(4):501–514, 2018. doi: 10.1109/TEVC.2017.2771451.
[16] C. Yang, J. Ding, Y. Jin, C. Wang, and T. Chai, "Multitasking multiobjective evolutionary operational indices optimization of beneficiation processes," *IEEE Trans. Automat. Sci. Eng.*, 16(3):1046–1057, 2019. doi: 10.1109/TASE.2018.2865593.
[17] A. Gupta, Y. Ong, and L. Feng, "Multifactorial evolution: Toward evolutionary multitasking," *IEEE Trans. Evol. Comput.*, 20(3):343–357, 2016. doi: 10.1109/TEVC.2015.2458037.
[18] L. Feng et al., "Explicit evolutionary multitasking for combinatorial optimization: A case study on capacitated vehicle routing problem," *IEEE Trans. Cybern.*, pp. 1–14, 2020. doi: 10.1109/TCYB.2019.2962865.
[19] A. T. W. Min, Y. Ong, A. Gupta, and C. Goh, "Multiproblem surrogates: Transfer evolutionary multiobjective optimization of computationally expensive problems," *IEEE Trans. Evol. Comput.*, 23(1):15–28, 2019. doi: 10.1109/TEVC.2017.2783441.
[20] H. Nazif and L. S. Lee, "Optimised crossover genetic algorithm for capacitated vehicle routing problem," *Appl. Math. Model.*, 36(5):2110–2117, 2012. doi: 10.1016/j.apm.2011.08.010.
[21] M. N. Omidvar, X. Li, Y. Mei, and X. Yao, "Cooperative co-evolution with differential grouping for large scale optimization," *IEEE Trans. Evol. Comput.*, 18(3):378–393, 2014. doi: 10.1109/TEVC.2013.2281543.
[22] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: NSGA-II," *IEEE Trans. Evol. Comput.*, 6(2):182C–197, 2002. doi: 10.1109/4235.996017.
[23] V. Maniscalco, S. G. Polito, and A. Intagliata, "Tree-based genetic algorithm with binary encoding for QoS routing," in *Proc. 7th Int. Conf. Innov. Mobile Internet Serv. Ubiquitous Comput.*, 2013, pp. 101–107. doi: 10.1109/IMIS.2013.26.
[24] U. K. Chakraborty and C. Z. Janikow, "An analysis of gray versus binary encoding in genetic search," *Inf. Sci.*, 156(3):253–269, 2003. doi: 10.1016/S0020-0255(03)00178-6.
[25] D. E. Goldberg. *Genetic Algorithms in Search, Optimization and Machine Learning*, 1st ed. Addison-Wesley, 1989.
[26] A. Gupta, Y. Ong, and L. Feng, "Insights on transfer optimization: Because experience is the best teacher," *IEEE Trans. Emerg. Topics Comput. Intell.*, 2(1):51–64, 2018. doi: 10.1109/TETCI.2017.2769104.
[27] S. Marsland, *Machine Learning: An Algorithmic Perspective*, 2nd ed. Chapman & Hall, 2014.
[28] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. The MIT Press, 2016.
[29] L. Shao, F. Zhu, and X. Li, "Transfer learning for visual categorization: A survey," *IEEE Trans. Neural Netw. Learn. Syst.*, 26(5):1019–1034, 2015. doi: 10.1109/TNNLS.2014.2330900.
[30] K. Weiss, T. M. Khoshgoftaar, and D. D. Wang, "A survey of transfer learning," *J. Big Data*, 3(9), 2016. doi: 10.1186/s40537-016-0043-6.
[31] Y. Jin and J. Branke, "Evolutionary optimization in uncertain environments-a survey," *IEEE Trans. Evol. Comput.*, 9(3):303–317, 2005. doi: 10.1109/TEVC.2005.846356.
[32] H. Fu, P. R. Lewis, B. Sendhoff, K. Tang, and X. Yao, "What are dynamic optimization problems?" in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2014, pp. 1550–1557. doi: 10.1109/CEC.2014.6900316.
[33] L. Zhou et al., "Solving dynamic vehicle routing problem via evolutionary search with learning capability," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2017, pp. 890–896. doi: 10.1109/CEC.2017.7969403.
[34] I. Hatzakis and D. Wallace, "Dynamic multi-objective optimization with evolutionary algorithms: A forward-looking approach," in *Proc. 8th Annu. Conf. Genetic Evol. Comput.*, 2006, pp. 1201–1208. doi: 10.1145/1143997.1144187.
[35] G. Ruan, L. L. Minku, S. Menzel, B. Sendhoff, and X. Yao, "When and how to transfer knowledge in dynamic multi-objective optimization," in *Proc. 2019 IEEE Symp. Series Comput. Intell. (SSCI)*, pp. 2034–2041. doi: 10.1109/SSCI44817.2019.9002815.
[36] M. Jiang, Z. Wang, L. Qiu, S. Guo, X. Gao, and K. C. Tan, "A fast dynamic evolutionary multiobjective algorithm via manifold transfer learning," *IEEE Trans. Cybern.*, early access, 2020. doi: 10.1109/TCYB.2020.2989465.
[37] M. Jiang, Z. Wang, H. Hong, and G. G. Yen, "Knee point based imbalanced transfer learning for dynamic multi-objective optimization," *IEEE Trans. Evol. Comput.*, early access, 2020. doi: 10.1109/TEVC.2020.3004027.
[38] J. Ding, C. Yang, Y. Jin, and T. Chai, "Generalized multitasking for evolutionary optimization of expensive problems," *IEEE Trans. Evol. Comput.*, 23(1):44–58, 2019. doi: 10.1109/TEVC.2017.2785351.
[39] K. K. Bali, A. Gupta, L. Feng, Y. S. Ong, and Tan Puay Siew, "Linearized domain adaptation in evolutionary multitasking," in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2017, pp. 1295–1302. doi: 10.1109/CEC.2017.7969454.

[40] J. Tang, Y. Chen, Z. Deng, Y. Xiang, and C. P. Joy, “A group-based approach to improve multifactorial evolutionary algorithm,” in *Proc. 27th Int. Joint Conf. Artif. Intell., IJCAI-18*, 2018, pp. 3870–3876. doi: 10.5555/3304222.3304307.
[41] A. Gupta, Y. Ong, L. Feng, and K. C. Tan, “Multiobjective multifactorial optimization in evolutionary multitasking,” *IEEE Trans. Cybern.*, 47(7):1652–1665, 2017. doi: 10.1109/TCYB.2016.2554622.
[42] R. Liaw and C. Ting, “Evolutionary many-tasking based on biocoenosis through symbiosis: A framework and benchmark problems,” in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2017, pp. 2266–2273. doi: 10.1109/CEC.2017.7969579.
[43] J. Zhong, L. Feng, W. Cai, and Y. Ong, “Multifactorial genetic programming for symbolic regression problems,” *IEEE Trans. Syst., Man, Cybern. Syst.*, pp. 1–14, 2018. doi: 10.1109/TSMC.2018.2853719.
[44] M. Gong, Z. Tang, H. Li, and J. Zhang, “Evolutionary multitasking with dynamic resource allocating strategy,” *IEEE Trans. Evol. Comput.*, 23(5):858–869, 2019. doi: 10.1109/TEVC.2019.2893614.
[45] H. Li, Y. Ong, M. Gong, and Z. Wang, “Evolutionary multitasking sparse reconstruction: Framework and case study,” *IEEE Trans. Evol. Comput.*, 23(5):733–747, 2019. doi: 10.1109/TEVC.2018.2881955.
[46] K. K. Bali, Y. Ong, A. Gupta, and P. S. Tan, “Multifactorial evolutionary algorithm with online transfer parameter estimation: MFEA-II,” *IEEE Trans. Evol. Comput.*, 24(1):69–83, 2020. doi: 10.1109/TEVC.2019.2906927.
[47] L. Feng et al., “Evolutionary multitasking via explicit autoencoding,” *IEEE Trans. Cybern.*, 49(9):3457–3470, 2019. doi: 10.1109/TCYB.2018.2845361.
[48] X. Ma, Q. Chen, Y. Yu, Y. Sun, L. Ma, and Z. Zhu, “A two-level transfer learning algorithm for evolutionary multitasking,” *Front. Neurosci.*, 13:1408, 2020. doi: 10.3389/fnins.2019.01408.
[49] G. Ochoa, J. Walker, M. Hyde, and T. Curtois, “Adaptive evolutionary algorithms and extensions to the hyflex hyper-heuristic framework, “ in *Proc. Parallel Problem Solving Nat. - PPSN XII*, 2012, pp. 418–427. doi: 10.1007/978-3-642-32964-7_42.
[50] E. Carrano, C. Fonseca, R. Takahashi, L. Pimenta, and O. Neto, “A preliminary comparison of tree encoding schemes for evolutionary algorithms,” in *Proc. IEEE Int, Conf. Syst., Man Cybern.*, 2007, pp. 1969–1974. doi: 10.1109/ICSMC.2007.4414042.
[51] S. J. Louis and J. McDonnell, “Learning with case-injected genetic algorithms,” *IEEE Trans. Evol. Comput.*, 8(4):316–328, 2004. doi: 10.1109/TEVC.2004.823466.
[52] P. Cunningham and B. Smyth, “Case-based reasoning in scheduling: Reusing solution components,” Tech. Rep., 1996.
[53] R. Santana, A. Mendiburu, and J. A. Lozano, “Structural transfer using EDAs: An application to multi-marker tagging SNP selection,” in *Proc. IEEE Congr. Evol. Comput.*, 2012, pp. 1–8. doi: 10.1109/CEC.2012.6252963.
[54] L. Feng, Y. Ong, M. Lim, and I. W. Tsang, “Memetic search with interdomain learning: A realization between CVRP and CARP,” *IEEE Trans. Evol. Comput.*, 19(5):644–658, 2015. doi: 10.1109/TEVC.2014.2362558.
[55] L. Feng, Y.S. Ong, A. H. Tan, and I. W. Tsang, “Memes as building blocks: a case study on evolutionary optimization + transfer learning for routing problems,” *Memetic Computing*, 7:159–180, 2015. doi: 10.1007/s12293-015-0166-x.
[56] A. Chaabani and L. B. Said, “Transfer of learning with the co-evolutionary decomposition-based algorithm-ii: A realization on the bi-level production-distribution planning system,” *Appl. Intell.*, 49(3):963–982, 2019. doi: 10.1007/s10489-018-1309-9.
[57] P. E. Mergos and A. Sextos, “Multi-objective optimum selection of ground motion records with genetic algorithms,” in *Proc. 16th Euro. Conf. Earthquake Eng.- 16ECEE*, 2018, pp. 18–21.
[58] C. Yang, J. Ding, Y. Jin, and T. Chai, “Off-line data-driven multi-objective optimization: Knowledge transfer between surrogates and generation of final solutions.” *IEEE Trans. Evol. Comput.*, 2019, pp 1–1. doi: 10.1109/TEVC.2019.2925959.
[59] L. Feng, Y. Ong, S. Jiang, and A. Gupta, “Autoencoding evolutionary search with learning across heterogeneous problems,” *IEEE Trans. Evol. Comput.*, 21(5):760–772, 2017. doi: 10.1109/TEVC.2017.2682274.
[60] J. Lin, H. Liu, K. C. Tan, and F. Gu, “An effective knowledge transfer approach for multiobjective multitasking optimization,” *IEEE Trans. Cybern.*, early access, pp. 1–11, 2020. doi: 10.1109/TCYB.2020.2969025.
[61] Z. Liang, W. Liang, X. Xu, and Z. Zhu, “A two stage adaptive knowledge transfer evolutionary multi-tasking based on population distribution for multi/many-objective optimization,” 2020.
[62] I. Steinwart and A. Christmann, *Support Vector Machines*, 1st ed. New York: Springer-Verlag, 2008.
[63] Y. Sun, B. Xue, M. Zhang, and G. G. Yen, “Evolving deep convolutional neural networks for image classification,” *IEEE Trans. Evol. Comput.*, vol. 24, no. 2, pp. 394–407, 2020. doi: 10.1109/TEVC.2019.2916183.
[64] Y. Sun, B. Xue, M. Zhang, G. G. Yen, and J. Lv, “Automatically designing CNN architectures using the genetic algorithm for image classification,” *IEEE Trans. Cybern.*, pp. 3840–3854, 2020. doi: 10.1109/TCYB.2020.2983860.
[65] A. L. Strehl and M. L. Littman, “Online linear regression and its application to model-based reinforcement learning,” in *Advances in Neural Information Processing Systems*, vol. 20, J. C. Platt, D. Koller, Y. Singer, and S. T. Roweis, Eds., 2008, pp. 1417–1424.
[66] B. Xue, M. Zhang, W. N. Browne, and X. Yao, “A survey on evolutionary computation approaches to feature selection,” *IEEE Trans. Evol. Comput.*, vol. 20, no. 4, pp. 606–626, 2016. doi: 10.1109/TEVC.2015.2504420.
[67] M. Iqbal, B. Xue, and M. Zhang, “Reusing extracted knowledge in genetic programming to solve complex texture image classification problems,” in *Proc. II, 20th Pacific-Asia Conf. Adv. Knowl. Discovery Data Mining*, 2016, vol. 9652, pp. 117–129. doi: 10.1007/978-3-319-31750-2_10.
[68] M. Iqbal, B. Xue, H. Al-Sahaf, and M. Zhang, “Cross-domain reuse of extracted knowledge in genetic programming for image classification,” *IEEE Trans. Evol. Comput.*, vol. 21, no. 4, pp. 569–587, 2017. doi: 10.1109/TEVC.2017.2657556.
[69] Q. Chen, B. Xue, and M. Zhang, “Generalisation and domain adaptation in gp with gradient descent for symbolic regression,” in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2015, pp. 1137–1144. doi: 10.1109/CEC.2015.7257017.
[70] M. Iqbal, W. N. Browne, and M. Zhang, “Reusing building blocks of extracted knowledge to solve complex, large-scale Boolean problems,” *IEEE Trans. Evol. Comput.*, vol. 18, no. 4, pp. 465–480, 2014. doi: 10.1109/TEVC.2013.2281537.
[71] T. Wei and J. Zhong, “A preliminary study of knowledge transfer in multi-classification using gene expression programming,” *Frontiers Neuroscience*, vol. 13, p. 1396, 2020. doi: 10.3389/fnins.2019.01396.
[72] J. B. Mouret and G. Maguire, “Quality diversity for multi-task optimization,” in *Proc. 2020 Genetic Evol. Comput. Conf.*, 2020, pp. 121–129. doi: 10.1145/3377930.3390203.
[73] A. D. Martinez, E. Osaba, J. D. Ser, and F. Herrera, “Simultaneously evolving deep reinforcement learning models using multifactorial optimization,” 2020.
[74] X. Li and X. Yao, “Cooperatively coevolving particle swarms for large scale optimization,” *IEEE Trans. Evol. Comput.*, vol. 16, no. 2, pp. 210–224, 2012. doi: 10.1109/TEVC.2011.2112662.
[75] Y. Mei, M. N. Omidvar, X. Li, and X. Yao, “A competitive divide-and-conquer algorithm for unconstrained large-scale black-box optimization,” *ACM Trans. Math. Softw.*, vol. 42, no. 2, pp. 1–24, 2016. doi: 10.1145/2791291.
[76] A. Kabán, J. Bootkrajang, and R. J. Durrant, “Toward large-scale continuous EDA: A random matrix theory perspective,” *Evol. Comput.*, vol. 24, no. 2, pp. 255–291, 2016.
[77] M. N. Omidvar, X. Li, and X. Yao, “Cooperative co-evolution with delta grouping for large scale non-separable function optimization,” in *Proc. IEEE Congr. Evol. Comput.*, 2010, pp. 1–8. doi: 10.1109/CEC.2010.5585979.
[78] Q. Yang, W. Chen, J. D. Deng, Y. Li, T. Gu, and J. Zhang, “A level-based learning swarm optimizer for large-scale optimization,” *IEEE Trans. Evol. Comput.*, vol. 22, no. 4, pp. 578–594, 2018. doi: 10.1109/TEVC.2017.2743016.
[79] B. Da, A. Gupta, Y. Ong, and L. Feng, “Evolutionary multitasking across single and multi-objective formulations for improved problem solving,” in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2016, pp. 1695–1701.
[80] D. Yu and L. Deng, *Automatic Speech Recognition: A Deep Learning Approach*. New York: Springer-Verlag, 2014.
[81] A. Karatzoglou and B. Hidasi, “Deep learning for recommender systems,” in *Proc. 11th ACM Conf. Recommender Syst.*, 2017, pp. 396–397.
[82] C. Zhang, P. Lim, A. K. Qin, and K. C. Tan, “Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics,” *IEEE Trans. Neural Netw. Learn. Syst.*, 28(10):2306–2318, 2017. doi: 10.1109/TNNLS.2016.2582798.
[83] J. Liu, M. Gong, Q. Miao, X. Wang, and H. Li. Structure learning for deep neural networks based on multiobjective optimization. *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 29, no. 6, pp. 2450–2463, 2018. doi: 10.1109/TNNLS.2017.2695223.
[84] A. V. Terekhov, G. Montone, and J. K. O’Regan, “Knowledge transfer in deep block-modular neural networks,” 2015, arXiv:abs/1908.08017.
[85] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, “How transferable are features in deep neural networks?” 2014, arXiv:abs/1411.1792.
[86] O. Fuentes S. Gutstein, and E. Freudenthal, “Knowledge transfer in deep convolutional neural nets,” *Int. J. Artif. Intell. Tools*, vol. 17, no. 3, pp. 555–567, 2008. doi: 10.1142/S0218213008004059.
[87] P. G. Ipeirotis, F. Provost, and J. Wang, “Quality management on Amazon Mechanical Turk,” in *Proc. ACM SIGKDD Workshop Human Comput.*, (HCOMP ‘10), 2010, pp. 64–67. doi: 10.1145/1837885.1837906.
[88] D. Delen, G. Walker, and A. Kadam, “Predicting breast cancer survivability: A comparison of three data mining methods,” *Artif. Intell. Med.*, vol. 34, no. 2, 2005. doi: 10.1016/j.artmed.2004.07.002.
[89] J. Cruz, P. Arias, D. Guimarans, D. Riera, and A. A. Juan, “Rich vehicle routing problem: Survey,” *ACM Comput. Surv.*, vol. 47, pp. 32:1–32:28, 2014. doi: 10.1145/2666003.
[90] L. Ge, J. Gao, H. Ngo, K. Li, and A. Zhang, “On handling negative transfer and imbalanced distributions in multiple source transfer learning,” *Stat. Anal. Data Min.*, pp. 254–271, 2014. doi: 10.1002/sam.11217.
[91] X. Yu, T. Liu, M. Gong, K. Zhang, K. Batmanghelich, and D. Tao, “Transfer learning with label noise,” 2017.
[92] B. Cao, S. J. Pan, Y. Zhang, D.-Y. Yeung, and Q. Yang, “Adaptive transfer learning,” 2010.
[93] J. Baxter, “A model of inductive bias learning,” *J. Artif. Intell. Res.*, vol. 12, no. 1, pp. 149–198, 2000. doi: 10.1613/jair.731.
[94] S. Ben-David and R. S. Borbely, “A notion of task relatedness yielding provable multiple-task learning guarantees,” *Mach. Learn.*, vol. 73, no. 3, pp. 273–287, Dec. 2008. doi: 10.1007/s10994-007-5043-5.
[95] L. Zhou, L. Feng, J. Zhong, Z. Zhu, B. Da, and Z. Wu, “A study of similarity measure between tasks for multifactorial evolutionary algorithm,” in *Proc. Genetic Evol. Comput. Conf. Companion*, 2018, pp. 229–230. doi: 10.1145/3205651.3205736.
[96] A. Gupta, Y. S. Ong, B. Da, L. Feng, and S. D. Handoko, “Landscape synergy in evolutionary multitasking,” in *Proc. IEEE Congr. Evol. Comput. (CEC)*, 2016, pp. 3076–3083. doi: 10.1109/CEC.2016.7744178.
[97] B. Da, A. Gupta, and Y. Ong, “Curbing negative influences online for seamless transfer evolutionary optimization,” *IEEE Trans. Cybern.*, vol. 49, no. 12, pp. 4365–4378, 2019. doi: 10.1109/TCYB.2018.2864345.
[98] K. K. Bali, A. Gupta, Y. Ong, and P. S. Tan, “Cognizant multitasking in multiobjective multifactorial evolution: MO-MFEA-II,” *IEEE Trans. Cybern.*, pp. 1–13, 2020.