Appl. Comput. Math., Vol.25, No.2, Special Issue, 2026, pp.270-298
DOI: 10.30546/1683-6154.25.2.2026.270

EFFICIENT BIOMEDICAL ONTOLOGY MATCHING FOR MEDICAL CYBER-PHYSICAL SYSTEMS THROUGH GENETIC ALGORITHM WITH SEMANTIC SEARCH AND ADAPTIVE SIMILARITY MEASURE

X. XUE[1,2,*], S.M. WAQAS[1], C.-M. CHEN[3], M. AMOON[4]

Abstract. In this paper we introduce a novel Genetic Algorithm with Semantic Search (GA-SS) to efficiently match biomedical ontologies. First, we use a new reference-free evaluation metric that combines an approximated F-measure estimator with a local consistency measure, enabling effective fitness evaluation without requiring gold-standard alignments. Second, we develop a semantic search mechanism that exploits structural regularities to efficiently refine promising solutions. Third, we design an adaptive similarity measure that integrates syntactic, linguistic, and taxonomic features and dynamically adjusts their weights based on their relevance to the matching context. The proposed GA-SS is compared with the state-of-the-art ontology matching methods on the Ontology Alignment Evaluation Initiative's Benchmark and Anatomy datasets, and the experimental results demonstrate that GA-SS is able to identify high-quality alignments while achieving superior matching efficiency, outperforming the state-of-the-art matching methods.

Keywords: Medical Cyber-Physical System, Biomedical Ontology Matching, Genetic Algorithm, Semantic Search, Adaptive Similarity Measure.

AMS Subject Classification: 68W50, 68T35, 92C50.

1. Introduction

Medical Cyber-Physical Systems (MCPS) [18] represent the tight integration of physical processes, biomedical data and networked computational resources in healthcare, enabling intelligent sensing, adaptive decision-making, predictive analytics, and real-time control. As MCPS become increasingly pervasive in diagnostics, monitoring, therapy, and rehabilitation [22], their effectiveness depends on the seamless semantic integration of heterogeneous biomedical data sources. The Semantic Web (SW) [8] provides a robust foundation for this integration by enriching the World Wide Web (WWW) with formalized knowledge representations. At the core of the SW are ontologies, which define structured vocabularies of domain-specific concepts and their interrelations [46], thereby enabling machines to consistently interpret, reason, and share biomedical knowledge. In practice, biomedical ontologies have been widely adopted to support applications ranging from Electronic Health Record (EHR) integration [3] to semantic-driven clinical decision support [26]. However, the decentralized development of ontologies and the subjectivity of biomedical knowledge modeling introduce inconsistencies in entity descriptions, giving rise to the ontology heterogeneity problem [23]. Addressing this challenge is particularly

[1]College of Artificial Intelligence, Yango University, Fuzhou, Fujian, 350015, China, e-mail: waqas@ygu.edu.cn
[2]Fujian Provincial Key Laboratory of Big Data Mining and Applications, Fujian University of Technology, Fuzhou, Fujian, 350118, China
[3]School of Artificial Intelligence, Nanjing University of Information Science and Technology, Nanjing, 211544, China, e-mail: chienmingchen@ieee.org
[4]Department of Computer Science and Engineering, College of Applied Studies, King Saud University, Riyadh, Saudi Arabia, e-mail: mamoon@ksu.edu.sa
*Corresponding author e-mail: jack8375@gmail.com
Manuscript received 06 October 2025.

critical for MCPS, where inconsistent interpretations of patient data, physiological parameters, and medical terminologies can directly compromise safety, accuracy, and responsiveness. To this end, ontology matching has emerged as a key research focus, which aims to identify the conceptual relationships between entities across ontologies. However, biomedical ontology matching remains highly challenging due to the massive scale of ontologies, their semantic richness, and the demand for efficient real-time reasoning in mission-critical MCPS applications.

This work focuses on resolving the one-to-one ontology matching problem between two ontologies, which is essentially a bipartite graph matching problem. Each ontology is represented as a graph, and the problem is to assign each node (or entity) in one graph (ontology) to at most one node in the other graph. Since the matching process aims to find the minimum edge dominating set [17], ontology matching is an NP-hard problem [27]. To find the optimal one-to-one alignment, exhaustive search requires a time complexity of $O(|\mathcal{O}_2| \times |\mathcal{O}_1|)$, where $|\mathcal{O}_1|$ and $|\mathcal{O}_2|$ are the numbers of entities in the ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$, respectively. This motivates the development of advanced computational intelligence techniques to ensure efficient and accurate alignments under the constraints of large-scale biomedical ontologies.

Existing approaches to ontology matching fall broadly into machine learning (ML)-based and heuristic categories. ML-based methods, including deep learning, often achieve high accuracy but require extensive training data, long training times, and significant computational resources, making them less suitable for dynamic or resource-constrained MCPS environments [31]. While heuristic methods such as rule-based [9] and greedy algorithms [7] offer faster runtimes and lower resource demands, they frequently compromise accuracy and robustness when applied to semantically heterogeneous biomedical ontologies. To address this issue, Genetic Algorithms (GA) have been widely used to achieve more effective ontology alignments [14], blending the speed of heuristic approaches like rule-based algorithms and greedy algorithms with improved alignment accuracy. Unlike rule-based methods, which rely on static logic, or greedy algorithms, which seek local optima, GAs perform a systematic search for global optima. By employing iterative evolution and population-based search, GAs overcome the limitations of initial configurations and local optima, enabling the discovery of more globally optimal ontology alignments.

Despite their success, existing GA-based ontology matching techniques still suffer from the following limitations that hinder their practical applicability especially for matching large-scale ontologies. First, they require experts to provide a reference alignment [10] or partial reference alignment [56] to evaluate the quality of the matching results and guide the search direction of the algorithm. When the scales of ontologies are large, it is impractical to require experts to match a large number of entities manually. In addition, a correct alignment should be consistent with the class hierarchical structure of ontologies [39]. Most existing GA-based matching methods ignore this local consistency principle [47], and thus yield unsatisfiable assertions in the integrated ontology. Another significant limitation of current GA-based matching methods in ontology alignment lies in their computationally intensive nature, primarily caused by the expansive solution space that impedes rapid convergence. They ignore the relationships between the genes in a chromosome, and evolve the population without taking into consideration the entities' semantic context. In this way, each source entity would have to compare with all the target entities independently, which leads to low search efficiency. Finally, similarity measures cannot effectively distinguish heterogeneous entities under different matching contexts. Existing methods typically use various similarity measures to independently calculate the similarity matrix for all potential entity pairs and then aggregate the similarity matrices. In general, each similarity matrix is treated as a black box, and their physical meanings are ignored. This can greatly reduce the confidence of the results and make it difficult to distinguish heterogeneous entities [54].

To overcome these challenges, we propose a Genetic Algorithm with Semantic Search and Adaptive Similarity Measure (GA-SS) for efficient biomedical ontology matching tailored to MCPS. The main contributions of the proposed approach are summarized as follows:

(1) To reduce dependency on reference alignments for match quality assessment, we propose two novel statistical evaluation metrics. The first proposed metric estimates the F-measure without requiring pre-defined reference alignments, while the second evaluates alignment quality through logical consistency analysis by identifying conflicting mappings. Based on these two metrics, a new fitness function is defined to guide the search process of the algorithm;
(2) To improve the search efficiency, a new semantic search method is developed by taking advantage of semantic relationships between genes and the local consistency principle. Specifically, it uses the biomedical ontology's concept hierarchical structure as the background knowledge to identify the semantically related entities under the hierarchical structure. Then, it conducts a local search within a much smaller neighborhood, which updates the semantically related entities only. By incorporating the semantic search into the GA framework, the search efficiency can be greatly improved;
(3) To improve the effectiveness of similarity measures, an adaptive entity similarity measure is proposed, which measures the similarity of two entities in terms of syntax, linguistic and taxonomy aspects simultaneously and dynamically adjusts the similarity score according to the differences of similarity measures and the biomedical entity correspondence's preference on them.

By integrating semantic awareness and adaptive evaluation into GA, the proposed GA-SS framework enhances both efficiency and reliability, making it particularly suited for ontology matching in MCPS, where timely and accurate semantic interoperability is indispensable.

The remainder of this paper is organized as follows. Section 2 introduces the definitions of ontology alignment and similarity measures. Section 3 reviews related work. Section 4 details the proposed GA-SS method. Section 5 reports experimental results and discussions. Section 6 concludes the paper and outlines the directions for future research.

2. Preliminaries

2.1. **Ontology Alignment.** A biomedical ontology is formally represented as a 5-tuple $\mathcal{O} = \langle \mathcal{C}, \mathcal{P}, \mathcal{I}, \Lambda, \Gamma \rangle$ [64], where $\mathcal{C}$ denotes the set of concepts (classes or categories of biomedical entities), $\mathcal{P}$ represents the properties that capture both attributes of concepts and relationships among them, and $\mathcal{I}$ corresponds to the set of instances, i.e., concrete biomedical entities such as patient cases, anatomical structures, or laboratory results. Properties in $\mathcal{P}$ may describe literal attributes (e.g., "age" of a patient) or relational links (e.g., "part-of" relations between anatomical components). Λ denotes the set of axioms defining logical constraints, and Γ contains metadata, such as comments or annotations, that clarify the intended meaning of classes. Collectively, $\mathcal{E} = \mathcal{C} \cup \mathcal{P} \cup \mathcal{I}$ defines the ontology entities.

Given two ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$, a correspondence between entities is defined as a 5-tuple $\langle id, e_1, e_2, r, c \rangle$ [68], where $e_1 \in \mathcal{E}_1$ and $e_2 \in \mathcal{E}_2$, id is a unique identifier of the correspondence, $r \in \{\equiv, \leq, \perp\}$ specifies the semantic relation between e_1 and e_2 (with $\equiv$ for equivalence, $\leq$ for subsumption, and $\perp$ for disjointness), and $c \in [0, 1]$ quantifies the confidence in this relation. Among these, identifying equivalence relations is of primary importance in biomedical ontology matching, as it enables interoperability across heterogeneous knowledge bases and supports the subsequent inference of more complex relationships. Accordingly, this work concentrates on discovering equivalence correspondences.

The biomedical ontology matching task is to derive an alignment $\mathcal{A}$, defined as a set of correspondences between entities in $\mathcal{O}_1$ and $\mathcal{O}_2$. Alignments may follow different cardinality constraints, including one-to-one, one-to-many, many-to-one, and many-to-many [37]. This study focuses on the one-to-one case, which forms the foundation for deriving other alignment types while ensuring clarity and consistency in the biomedical entity mappings. Beyond the ontologies themselves, additional inputs may be available: (i) a historical alignment $\mathcal{A}'$, which provides

prior knowledge to guide the current matching; (ii) a parameter set p, including thresholds for similarity values, weighting schemes for syntactic, semantic, or taxonomic similarity measures, and algorithmic hyperparameters; and (iii) external resources r, such as domain-specific lexicons or biomedical knowledge bases (e.g., UMLS, MeSH), which provide auxiliary semantic evidence to improve alignment reliability.

Fig.1 illustrates the general workflow of ontology matching: (1) Computing similarity values for candidate entity pairs, (2) identifying semantically similar entities according to selected similarity measures and thresholds, and (3) validating and refining the resulting alignments. The inputs include ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$, optional historical alignment $\mathcal{A}'$, parameter set p, and external resources r, while the output is the final alignment $\mathcal{A}$. Therefore, the quality of an ontology matching system depends critically on three factors: (1) How entity similarity is measured, (2) how correct correspondences are efficiently identified in the vast candidate space, and (3) how incorrect correspondences are effectively filtered to preserve alignment accuracy and consistency.

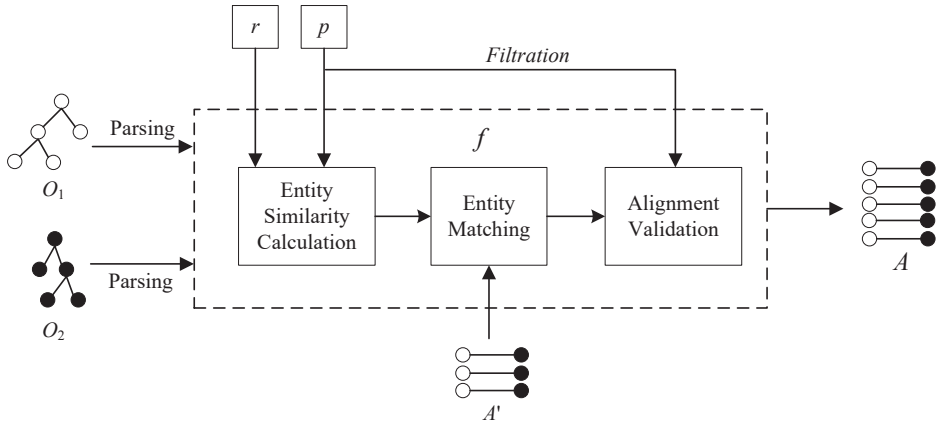


FIGURE 1. A typical biomedical ontology matching process.

In this work, our focus is on establishing one-to-one alignment, which is the most fundamental task in ontology matching. Other types of alignment can be derived, either directly or indirectly, from this one-to-one alignment. In addition to the given ontologies $\mathcal{O}_1$ and $\mathcal{O}_2$, our approach may incorporate additional information, such as an optional historical alignment $\mathcal{A}'$, a specific parameter set p, and external resources r (e.g., background knowledge bases to differentiate domain-specific terms). The quality of an alignment $\mathcal{A}$ is commonly evaluated against a reference alignment $\mathcal{R}$, which is not observed during the ontology matching process. The goal is to make the alignment $\mathcal{A}$ as close to the reference alignment $\mathcal{R}$ as possible. Specifically, the recall, precision and F-measure are calculated as follows [36]:

$$\begin{aligned}
\texttt{recall}(\mathcal{A};\mathcal{R}) &= \frac{|\mathcal{R}\cap\mathcal{A}|}{|\mathcal{R}|},\\
\texttt{precision}(\mathcal{A};\mathcal{R}) &= \frac{|\mathcal{R}\cap\mathcal{A}|}{|\mathcal{A}|},\\
\texttt{f}(\mathcal{A};\mathcal{R}) &= \frac{2\times\texttt{recall}(\mathcal{A};\mathcal{R})\times\texttt{precision}(\mathcal{A};\mathcal{R})}{\texttt{recall}(\mathcal{A};\mathcal{R})+\texttt{precision}(\mathcal{A};\mathcal{R})},
\end{aligned} \tag{1}$$

where $|\cdot|$ is the cardinality (number of correspondences) of the alignment. $|\mathcal{R}\cap\mathcal{A}|$ indicates the number of common correspondences in $\mathcal{A}$ and $\mathcal{R}$, which is also the number of correct correspondences in $\mathcal{A}$. If the recall reaches 1, it indicates that every correspondence in the reference set

$\mathcal{R}$ has been successfully retrieved in the alignment set $\mathcal{A}$. When precision equals 1, all elements identified in $\mathcal{A}$ are guaranteed to belong to $\mathcal{R}$. A perfect F-measure of 1 occurs only when $\mathcal{A}$ and $\mathcal{R}$ coincide completely, containing identical correspondences.

Overall, the ontology matching problem is to find an alignment $\mathcal{A}^*$ that maximizes the F-measure, as shown in Eqn. (1), and there are two challenges: (1) How to evaluate the fitness value of an individual because the reference alignment $\mathcal{R}$ is often unknown in practice, and (2) how to efficiently find the optimal one-to-one alignment from the huge search region. To tackle these issues, we propose the following: (1) A new evaluation metric $\texttt{fit}(\cdot)$ (see also Section 4.3) to approximate its F-measure value; and (2) a SS strategy to improve the algorithm's search efficiency (see also Section 4.5).

2.2. **Similarity Measures.** In ontology matching, the fundamental objective is to evaluate how closely entities from distinct ontologies correspond, thereby enabling the discovery and alignment of equivalent pairs in the ontology. Methods for estimating such similarity are generally organized into three major categories [64], each emphasizing different perspectives of entity comparison: syntactic techniques that rely on surface-level character patterns, linguistic approaches that exploit semantic and lexical information, and taxonomy-based measures that incorporate hierarchical structural relations among entities.

2.2.1. *Syntactic-based Similarity Measure.* Syntactic similarity measures assess how closely two entities match by computing their string or edit distances. In this work, we employ two widely used approaches: the String Metric for Ontology Alignment (SMOA) [48] and the Levenshtein Distance (LD) [44]. For a given pair of entities e_1 and e_2, SMOA and LD values are computed as follows:

$$\texttt{SMOA}(e_1, e_2) = \texttt{comm}(e_1, e_2) - \texttt{diff}(e_1, e_2) + \texttt{imp}(e_1, e_2),$$

where $\texttt{comm}(e_1, e_2)$ denotes the length of the overlapping substrings between e_1 and e_2, and $\texttt{diff}(e_1, e_2)$ represents the unmatched portion of the two strings. The adjustment function $\texttt{imp}(e_1, e_2)$ refines the similarity estimation by leveraging prefix-based modifications, which extend Winkler's approach [57] through the explicit consideration of initial substring matches.

$$\texttt{LD}(e_1, e_2) = \max\left\{0, \frac{\min(|e_1|, |e_2|) - d(e_1, e_2)}{\min(|e_1|, |e_2|)}\right\},$$

where $|e_1|$ and $|e_2|$ are the number of characters in the e_1 and e_2 strings, respectively. $d(e_1, e_2)$ is the number of characters that need to be changed from the e_1 string to the e_2 string.

For considering both SMOA and LD simultaneously, the aggregated syntactic-based similarity measure is defined by averaging over SMOA and LD as follows [19]:

$$\texttt{sim}_s(e_1, e_2) = \frac{1}{2}(\texttt{SMOA}(e_1, e_2) + \texttt{LD}(e_1, e_2)).$$

2.2.2. *Linguistic-based Similarity Measure.* A linguistic similarity measure evaluates the semantic relationship between the names of two entities, often leveraging synonym resources such as the widely used WordNet electronic lexical database [30]. For two entities e_1 and e_2 with names w_1 and w_2, the linguistic similarity $\texttt{sim}_l(e_1, e_2)$ is defined as follows [25]:

$$\texttt{sim}_l(e_1, e_2) = \begin{cases} 1, & \text{if } w_1 \text{ and } w_2 \text{ are synonyms,} \\ 0.5, & \text{if } w_1 \ (w_2) \text{ is } w_2 \ (w_1)\text{'s hypernym,} \\ 0, & \text{otherwise.} \end{cases}$$

2.2.3. *Taxonomy-based Similarity Measure.* Taxonomy-based similarity measures assess the similarity of two entities by comparing their respective neighborhoods, typically by leveraging the hierarchical structure of the ontology. Given two entities e_1 and e_2, below are two taxonomy-based similarity measures between them. The first one is based on the set of direct children (sub concepts) $\mathcal{C}_1$ and $\mathcal{C}_2$ in the topology graph, which is defined as follows [34]:

$$\mathtt{sim}_{conc}(e_1, e_2) = \frac{\sum_{k=1}^{2} \sum_{c \in \mathcal{C}_k} \max_{c' \in \mathcal{C}_{3-k}} \mathtt{sim}_{sl}(c, c')}{|\mathcal{C}_1| + |\mathcal{C}_2|},$$

where the function $|\cdot|$ indicates the number of elements in the set, and the similarity measure $\mathtt{sim}_{\mathtt{sl}}(\cdot)$ is defined in Eqn. 5. The above concept-based similarity measure stands for the average similarity of all the direct children of e_1 and e_2 with their most similar entities in the other ontology.

The second taxonomy-based similarity measure considers the common instances shared by the two entities (concepts). Specifically, the instance-based similarity measure is defined as follows [2]:

$$\mathtt{sim}_{inst}(e_1, e_2) = \frac{|\mathcal{I}(e_1) \cap \mathcal{I}(e_2)|}{\max\{|\mathcal{I}(e_1)|, |\mathcal{I}(e_2)|\}},$$

where $\mathcal{I}(e_1)$ and $\mathcal{I}(e_2)$ are the corresponding set of instances to e_1 and e_2, respectively. $|\mathcal{I}(e_1) \cap \mathcal{I}(e_2)|$ is the number of instances belonging to both entities. If the two entities have no common instance, then they are least similar, and $\mathtt{sim}_{inst}(e_1, e_2) = 0$. If the two entities share exactly the same set of instances, then they are considered as equivalent entities, and their instance-based similarity reaches the maximum of 1.

3. Related work

In the past decades, many ontology matching techniques have been proposed to find the semantically similar entities that come from two heterogeneous ontologies. These methodologies are primarily classified into two main categories: learning-based methods and heuristic-based approaches. Learning-based matching techniques use different ML techniques, such as the Bayesian Network (BN) [67], Decision Tree (DT) [40], Support Vector Machine (SVM) [29], Random Forest [51], and Convolutional Neural Network (CNN) [28], to learn knowledge from the sample dataset provided by human experts or from the external knowledge base [51]. The learned rules or knowledge are then used to find similar ontology entities. To alleviate the reliance of ML-based methods on extensive but often scarce labeled training data, recent advancements have used weak supervision sources that generate labels from a large corpus of unlabeled data [15] or automatically construct the training dataset based on external knowledge bases [24], thereby facilitating the training of ML models.

Compared with learning-based methods, heuristic-based approaches do not need to build the training dataset and execute the complex learning process, which could be more efficient. The existing heuristic-based approaches can be further classified into rule-based approaches and meta-heuristic approaches. The rule-based matching techniques use the knowledge or rules provided by human experts to prune the search space [45, 70], select [35], combine [59] and tune [20] different similarity measures, and extract the alignment [5]. However, their reliance on expert knowledge to identify similar entity pairs limits their effectiveness across diverse matching scenarios, particularly when handling complex heterogeneity. Therefore, in recent years, meta-heuristics such as Evolutionary Algorithms (EAs) [38] have attracted extensive attention in the ontology matching community, which aim to optimize the alignment by meta-heuristic algorithms.

EA-based ontology matching techniques can be divided into two categories, i.e., the ontology meta-matching method and the ontology entity matching approach [62]. The first category

models ontology matching as an aggregation problem of various similarity measures, aiming to discover the most effective method for combining these measures. However, these methods have to maintain the similarity matrices during the optimization of aggregation parameters, resulting in high computational complexity [60]. To overcome this drawback, the second category directly searches for the optimal ontology alignment based on iterative improvement frameworks. The authors of the paper [53] first defined the intentional and extensional features on the entities, and then used GA to find a near-optimal alignment. In [42], the authors also considered ontology matching as the process of finding an ideal mapping from a set of existing correspondences, and they used GA to directly extract the final alignment. In [1], the authors proposed to use GA with stochastic hill climbing algorithm for finding the optimal alignment, which is the milestone of EA-based entity matching technique. In [4], the authors argued that the instance information of two ontologies can more precisely measure the similarity between their corresponding concepts. They proposed to first find the shared instances of the two ontologies, and then used hybrid GA to find the similar concepts based on the instance alignment. However, the instance-based similarity measure has the following limitations: (1) the scale of instances is always huge, and matching instances require much larger computational cost than directly matching classes; (2) it is rare that two concepts own the same or similar instance sets, and the lack of common instances between concepts yield low accuracy of matching results. To further improve the performance of GA-based techniques, the authors of the paper [65] proposed a problem-specific Interactive GA (IGA), which involves users into the evolutionary process to identify the problematic correspondences and guide the subsequent search direction. However, the manual identification is a time consuming and strenuous process, and human experts would unavoidably make mistakes when dealing with huge quantity of information coming from the corresponding ontology entities. Recently, they further developed a novel kernel principal component analysis and evolutionary algorithm (kPCA-EA) [66], which employed cosine similarity-driven kPCA to capture complex Similarity Feature (SF) relationships and automatically constructs SF for OM, thereby guiding the algorithm to achieve higher-level ontology alignment results. However, the exclusive reliance on cosine similarity for measuring SF correlations may result in suboptimal performance across diverse matching scenarios where alternative similarity metrics could be more appropriate. Although EA-based entity matching techniques are more efficient than meta-matching approaches in terms of computational cost, they suffer from the slow convergence, and are easy to get stuck in the local optima due to the large search space of ontology entity matching problem.

Although the existing GA-based ontology matching methods can achieve reasonably competitive results, the following two drawbacks hamper their applications in the practical scenarios: (1) dependence on manually defined reference alignments for evaluating fitness, which is often impractical in real-world settings, and (2) difficulty maintaining efficiency when handling heterogeneous entities and large-scale ontologies, leading to slow convergence. To address these limitations, this paper introduces a GA-SS, which integrates semantic guidance, local consistency awareness, and dynamic similarity adaptation to achieve more efficient and reliable ontology alignment.

4. Genetic algorithm with semantic search for ontology matching

4.1. **Algorithm Overview.** The framework of GA-SS follows an evolutionary optimization process that explores a large solution space while explicitly considering semantic relationships among ontology entities. This design enables the discovery of accurate and consistent ontology alignments. The term semantic search in GA-SS refers to a guided local search strategy that exploits the hierarchical and semantic structure of biomedical ontologies to refine promising candidate alignments. It restricts the search to semantically related entities rather than performing global comparisons, thereby improving both accuracy and efficiency. The proposed GA-SS pseudo-code is shown in Algorithm 1.

Algorithm 1: Genetic Algorithm with Semantic Search

Input: Population $\mathcal{P}$, Crossover Rate $rate_c$, Mutation Rate $rate_m$, Elite Individuals $\mathcal{I}_E$
Output: $\mathcal{I}_E$

Initialization Population $\mathcal{P}$;
$\mathcal{I}_E \leftarrow \emptyset$;
Evaluation($\mathcal{P}$);
Generation $gen \leftarrow 0$;
while $gen < MaximumGen_m$ **do**
    Parent Selection $\mathcal{P}'$ with $\mathcal{P}$;
    /* Breeding (genetic operators) */
    $\mathcal{P}'' \leftarrow crossover(P', rate_c)$;
    $\mathcal{P}'' \leftarrow mutation(P', rate_m)$;
    Evaluation($\mathcal{P}''$);
    $Update(\mathcal{I}_E)$;
    Semantic Search($\mathcal{I}_E$);
    $\mathcal{P} \leftarrow environment_selection()$;
    $gen \leftarrow gen + 1$;
end
return $\mathcal{I}_E$;

First, the population $\mathcal{P}$ is randomly initialized and each initial individual is evaluated through the new Fitness Function. During each iteration, the evolutionary process begins with tournament-based parent selection $\mathcal{P}'$, which favors individuals with higher fitness values while maintaining population diversity. The selected parents undergo crossover and mutation operations at rates $rate_c$ and $rate_m$, respectively, to generate new offspring. Each offspring is then evaluated and compared with the current best (elite) individual $\mathcal{I}_E$. After evaluation, the semantic search procedure is applied to the elite individual. This process explores a restricted neighborhood defined by the ontology's class hierarchy to correct or refine semantically inconsistent alignments, thereby maintaining local consistency. The population is then updated through environmental selection. The current and offspring populations are merged, sorted in descending order of fitness, and the top 50% of individuals are retained using roulette-wheel selection. This ensures a balance between exploitation of high-quality solutions and exploration of new regions of the search space. The evolution continues until either an optimal fitness value of 1.0 is achieved or the maximum number of generations is reached.

Table 1 provides a concise summary of these parameters, their conceptual purposes, and related algorithmic components. The specific values used in experiments are detailed in Section 5.1.

TABLE 1. Summary of GA-SS parameters and their conceptual roles.

Parameter	Conceptual Role	Associated Component
$rate_c$	Controls Offspring Generation via Crossover	Evolutionary Operator
$rate_m$	Maintains Population Diversity via Mutation	Evolutionary Operator
α	Balances Fitness between Accuracy and Consistency	Fitness Function
θ, ϕ	Weight syntactic, linguistic, and taxonomic similarities	Adaptive Similarity Measure
ρ	Threshold for anchor correspondences in semantic search	Semantic Search

Compared with the standard GA framework, the proposed GA-SS contains two novel components, i.e., the Fitness Function and Semantic Search, which will be described in more detail in Sections 4.3 and 4.5.

4.2. **Individual Representation.** In GA-SS, each individual (alignment) is represented as a $|\mathcal{O}_1|$-dimensional vector $X = [x_1, \ldots, x_{|\mathcal{O}_1|}]$, where $x_i \in \{0, 1, \ldots, |\mathcal{O}_2|\}$ is an entity index in $\mathcal{O}_2$, indicating the entity in $\mathcal{O}_2$ that is matched to entity i in $\mathcal{O}_1$. Note that $x_i = 0$ means that entity i in $\mathcal{O}_1$ is matched with no entity in $\mathcal{O}_2$.

Fig.2 shows an example of the individual representation. As shown in the figure, there are 8 entities in ontology $\mathcal{O}_1$ and 9 entities in ontology $\mathcal{O}_2$. Therefore, the vector has 8 dimensions, each of which can take values from 0 to 9. In the example individual, $x_2 = 4$ means that the 2^{nd} entity "oropharynx epithelium" in $\mathcal{O}_1$ is mapped with the 4^{th} entity "Oropharynx_Epithelium" in $\mathcal{O}_2$, and $x_5 = 0$ means that the 5^{th} entity "gonad" in $\mathcal{O}_1$ is mapped with no entity.

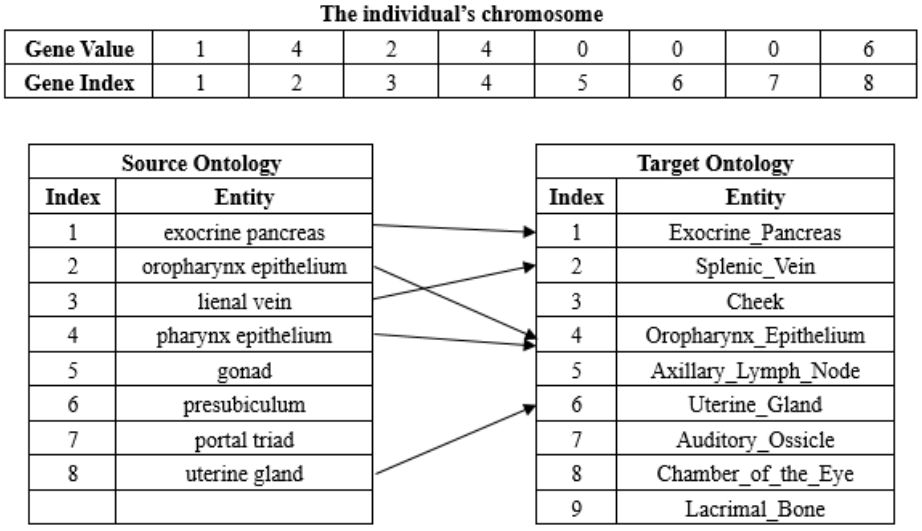


FIGURE 2. An example of the individual (alignment) representation.

Note that we focus on one-to-one alignment in this work, i.e., each entity in $\mathcal{O}_1$ is mapped with at most one entity in $\mathcal{O}_2$. To this end, each entity in $\mathcal{O}_2$ must occur at most once in the vector. Here, we handle this constraint implicitly. That is, we ignore the constraint in the individual representation (i.e., the entities in $\mathcal{O}_2$ can occur any number of times), but design the fitness function in a way that the one-to-one alignments tend to have better fitness values.

4.3. **Fitness Function.** The design of fitness function is critical to guide the search towards the target alignment. In other words, the fitness function should be designed in a way that the (unseen) target alignment has the best fitness, and an alignment that is more similar to the target alignment tends to have a better fitness. Since the target (reference) alignment is unknown in advance, we cannot directly utilize the F-measure (shown in Eqn. (1)) as the fitness function. Therefore, the challenge in designing the fitness function is to approach the real F-measure as closely as possible.

To address this challenge, we propose a new fitness function that approximates the F-measure based on the following two aspects: (1) approximating the raw F-measure by estimating the precision and recall, and (2) locality that reflects the hierarchical consistency of the alignment. Specifically, the fitness function of an alignment $\mathcal{A}$ is defined as follows:

$$\texttt{fit}(\mathcal{A}) = \alpha \cdot \texttt{f}'(\mathcal{A}) + (1 - \alpha) \cdot \texttt{loc}(\mathcal{A}).$$

where $\texttt{f}'(\mathcal{A})$ is the approximated F-measure, and $\texttt{loc}(\mathcal{A})$ is the local consistency measure. The parameter $\alpha \in [0, 1]$ controls the balance between these two components.

4.3.1. *Approximated F-measure.* The approximated F-measure $\mathtt{f}'(\mathcal{A})$ is defined as follows:

$$\mathtt{f}'(\mathcal{A}) = \frac{2 \cdot \mathtt{r}'(\mathcal{A}) \cdot \mathtt{p}'(\mathcal{A})}{\mathtt{r}'(\mathcal{A}) + \mathtt{p}'(\mathcal{A})}, \tag{2}$$

$$\mathtt{r}'(\mathcal{A}) = \frac{|\mathcal{A}|}{\max\{|\mathcal{O}_1|, |\mathcal{O}_2|\}}, \tag{3}$$

$$\mathtt{p}'(\mathcal{A}) = \frac{1}{|\mathcal{A}|} \sum_{(e_1,e_2)\in\mathcal{A}} \mathtt{sim}(e_1, e_2). \tag{4}$$

Eqn. (3) estimates the recall as the number of entity correspondences in $\mathcal{A}$ divided by the maximal possible number of one-to-one correspondences in an alignment. Obviously, increased values of $\mathtt{r}'(\mathcal{A})$ correlate strongly with greater coverage of target alignment correspondences, consequently yielding enhanced recall metrics. Eqn. (4) estimates the precision based on the similarities between the matched entities in $\mathcal{A}$. Intuitively, if two entities have higher similarity, then they are more likely to be matched in the target alignment. Thus, by maximizing $\mathtt{p}'(\mathcal{A})$ in Eqn. (4), we tend to select only the high-similarity correspondences that belong to the target alignment in $\mathcal{A}$, leading to higher precision.

4.3.2. *Adaptive Similarity Measure.* The effectiveness of ontology alignment depends heavily on the choice of similarity measure in Eqn. 4, as it determines how accurately semantic relationships between entities are captured. Ontologies exhibit unique lexical and structural characteristics that make this task challenging. First, many biomedical ontologies use structured naming conventions or standardized abbreviations, which create high syntactic regularity but also cause synonym mismatches across domains. Second, ontologies are typically organized into deep hierarchical structures that contain rich taxonomic information, while often lacking sufficient instance data for reliable instance-based reasoning. Third, frequent use of technical terminology and alphanumeric identifiers reduces the reliability of purely linguistic similarity measures. These challenges indicate that no single similarity dimension-syntactic, linguistic, or taxonomicis-sufficient to capture semantic equivalence among entities [11]. To address this, we define an **adaptive similarity measure**, which dynamically integrates syntactic, linguistic, and taxonomy-based evidence according to the characteristics of each entity pair. This adaptive mechanism enables more accurate and robust estimation of similarity across heterogeneous biomedical ontologies.

Although both syntax-based and linguistic-based similarity measures calculate the similarity value based on entity names, they consider different aspects. It is possible for entities to have very high syntactic similarity but very low linguistic similarity, and vice versa. Thus, it is important to consider both the syntactic-based and linguistic-based similarity measures to have a comprehensive estimation of the similarity between entities. The most commonly used aggregation methods are based on the maximum, minimum and average strategies [21]. The maximum (minimum) strategy aggregates the two kinds of similarity values by taking the maximum (minimum) between them. The maximum strategy ensures the completeness of the alignment, whereas the minimum strategy tends to increase correctness. To consider both completeness and correctness, we aggregate these two similarity measures by calculating their mean value, which is defined as follows:

$$\mathtt{sim}_{sl}(e_1, e_2) = \frac{1}{2}(\mathtt{sim}_s(e_1, e_2) + \mathtt{sim}_l(e_1, e_2)). \tag{5}$$

On the other hand, the neighbor concept-based and instance-based similarity measures reflect the taxonomy-based similarity from different aspects, i.e., concept-level and instance-level, respectively. In practice, the number of instances is often limited, leading to inaccurate estimation of instance-based similarity measures. To increase the estimation accuracy of the taxonomy-based similarity, we use the maximum strategy to aggregate the neighbor concept-based and

instance-based similarity measures, which is defined as follows:

$$\mathtt{sim}_t(e_1, e_2) = \max\{\mathtt{sim}_{conc}(e_1, e_2), \mathtt{sim}_{inst}(e_1, e_2)\}.$$

Finally, to aggregate the $\mathtt{sim}_{sl}(\dot{)}$ and $\mathtt{sim}_t(\dot{)}$ that consider two entities' similarity from totally different perspectives, the commonly used aggregating strategies, such as the maximize, minimize and average strategies, have shown to be ineffective in some cases, since they neglect the effect brought by the differences between different types of similarity measures and the preference of entity pairs for different similarity measures [1]. To address these constraints, we introduce a new hybrid similarity metric for entities e_1 and e_2, formally defined as:

$$\mathtt{sim} = \frac{1}{2}\left(\mathtt{sim}'_{sl} + \mathtt{sim}'_t\right), \tag{6}$$

where

$$\mathtt{sim}'_{sl} = \mathtt{sim}_{sl} + \begin{cases} |\mathtt{sim}_{sl} - \mathtt{sim}_t| \times \theta, & \text{if } \mathtt{sim}_{sl} > \phi, \\ -|\mathtt{sim}_{sl} - \mathtt{sim}_t| \times \theta, & \text{otherwise,} \end{cases} \tag{7}$$

$$\mathtt{sim}'_t = \mathtt{sim}_t + \begin{cases} |\mathtt{sim}_{sl} - \mathtt{sim}_t| \times \theta, & \text{if } \mathtt{sim}_t > \phi, \\ -|\mathtt{sim}_{sl} - \mathtt{sim}_t| \times \theta, & \text{otherwise.} \end{cases} \tag{8}$$

The new similarity measure is an average of two aspects: (1) syntactic and linguistic $\mathtt{sim}'_{sl}$ and (2) taxonomy $\mathtt{sim}'_t$. The former is related to the entity itself, while the latter focuses more on its neighbors in the ontology.

The syntactic and linguistic component is transformed from the aggregated similarity measure $\mathtt{sim}_{sl}$ given in Eqn. (5). We compare it with a threshold ϕ, then increase it if it is larger than the threshold (the two entities are similar in this aspect), while decrease it otherwise (they are dissimilar). The increased (decreased) amount $|\mathtt{sim}_{sl} - \mathtt{sim}_t| \times \theta$ depends on their similarity in these two aspects, where θ is a user-defined parameter. This way, the following scenarios is expected to occur:

- If the two entities are similar in the syntactic and linguistic aspect ($\mathtt{sim}_{sl} > \phi$) and the syntactic and linguistic similarity $\mathtt{sim}_{sl}$ is much higher than the taxonomy similarity $\mathtt{sim}_t$, then $\mathtt{sim}'_{sl}$ is much larger than $\mathtt{sim}_{sl}$;
- If the two entities are dissimilar in the syntactic and linguistic aspect ($\mathtt{sim}_{sl} \leq \phi$) and $\mathtt{sim}_{sl}$ is much lower than $\mathtt{sim}_t$, then $\mathtt{sim}'_{sl}$ is much smaller than $\mathtt{sim}_{sl}$;
- If $\mathtt{sim}_{sl}$ is similar to $\mathtt{sim}_t$, then $\mathtt{sim}'_{sl}$ is similar to $\mathtt{sim}_{sl}$.

We transform the taxonomy-based similarity measure in the same way, as shown in Eqn. (8). The proposed hybrid similarity measure aims to overcome the limitations of traditional aggregation methods by taking into account the differences between different types of similarity measures and the preference of entity pairs for each measure. To achieve this, we introduce two thresholds, θ and ϕ, to adjust the final similarity value. The weight factor θ determines the degree to which the difference between the two similarity measures influences the final score, while the threshold value ϕ controls the level of similarity required for the difference to impact the final score. By incorporating these two parameters, our proposed hybrid similarity measure is able to capture the distinctions between different types of similarity measures and the preferences of entity pairs for each measure, thus providing a more comprehensive and accurate estimation of entity similarity.

4.3.3. *Local Consistency Measure.* Beyond the reference-free F-measure estimator, a local consistency term is incorporated to maintain structural coherence by exploiting the hierarchical organization of ontologies. Ontologies are typically structured as taxonomies that describe functional, spatial, or part-whole relationships among domain entities. These hierarchies are semantically rich, with parent-child relations encoding critical domain knowledge such as containment, dependency, and specialization. If the established mappings violate these hierarchical dependencies, the resulting alignments become semantically inconsistent and may distort the

meaning of the integrated ontology. To prevent such inconsistencies, a local consistency constraint is introduced, requiring that whenever a source entity has a parentchild relation with another source entity, their corresponding target entities preserve the same relational structure. By enforcing this constraint, the alignment process favors structurally consistent mappings that more accurately capture the underlying semantics of complex biomedical or engineering systems.

Specifically, given two entity correspondences (e_1, e_2) and (e'_1, e'_2), if e_1 is the super concept (or sub concept) of e'_1, then the local consistency requires that e'_2 must be the super concept (or sub concept) of e_2. The local consistency measure of an alignment $\mathcal{A}$ is calculated as follows:

(1) Sort the correspondences in $\mathcal{A}$ in the descending order of their similarity values, initialise an empty set $\mathcal{A}'$;
(2) Add the sorted correspondences one by one into $\mathcal{A}'$, if it does not violate the local consistency with the correspondences in $\mathcal{A}'$.
(3) Calculate $\texttt{loc}(\mathcal{A}) = \frac{|\mathcal{A}'|}{|\mathcal{A}|}$.

Because ontologies tend to have well-behaved, tree-like hierarchies, this greedy selection yields a near-maximal consistent subset at negligible cost, allowing GA-SS to favor alignments that preserve engineering semantics while still exploring a broad search space.

4.4. **Genetic Operators.** We employ a two-point crossover operator [52] and a bitwise mutation operator [32] to generate offspring solutions during the evolutionary process. These genetic operators are applied to the vector-based encoding of individuals, where each gene represents a selected entity index or structural component from the target ontology $\mathcal{O}_2$.

The two-point crossover operator first selects two parent chromosomes from the mating pool according to a predefined selection strategy. Two crossover points are then randomly chosen along the chromosome length to define a contiguous segment, and the genetic material within this segment is exchanged between the parents to produce two offspring. This operation preserves useful partial structures, i.e., sub-alignments, while promoting exploration through recombination of different parental components.

After crossover, a bitwise mutation operator introduces random variation into each offspring. Each gene is mutated with a fixed probability p_m by replacing its current value with a randomly selected index from the set $\{0, 1, \ldots, |\mathcal{O}_2|\}$, where $|\mathcal{O}_2|$ denotes the number of available entities in the target ontology. This stochastic perturbation maintains population diversity and prevents premature convergence.

Together, the crossover and mutation operators facilitate a balanced search by combining exploitation of high-quality partial solutions with exploration of novel regions in the solution space. Their synergistic effect is crucial for constructing accurate and diverse similarity feature combinations or candidate mappings throughout the evolution.

4.5. **Semantic Search.** Ontologies possess several inherent characteristics that our proposed semantic search mechanism is designed to leverage. First, they typically exhibit deep but well-organized hierarchical structures that define clear parent-child relationships among concepts. Second, entities are often labeled using standardized terminologies or domain-specific descriptors, leading to consistent naming patterns and strong lexical coherence among semantically related entities. Third, ontologies adhere to local structural constraints, where sub-concepts inherit key semantic properties such as spatial, functional, or compositional attributes from their parent concepts. As a result, mappings that preserve such local consistency are more likely to be semantically valid. These characteristics collectively motivate the design of a locality-aware semantic search mechanism that refines ontology alignments efficiently by focusing on semantically related regions rather than performing exhaustive pairwise comparisons.

The pseudo-code of designed semantic search is presented in Algorithm 2. Given an individual X, we first initializes a set of anchor correspondences, where an anchor correspondence is a concept correspondence (c_i, c_{x_i}) with high similarity, where $\texttt{sim}(i, x_i) > \rho$ and $(i, x_i) \in X$. After that, for each anchor correspondence (c_i, c_{x_i}), we examine the possible correspondence (c_j, c_k),

where c_j and c_k are c_i and c_{x_i}'s sub-concepts in two ontologies, respectively. In particular, in an ontology hierarchical structure, a sub concept is a more specific or detailed version of a broader concept. For example, in an ontology about animals, "dog" would be a sub-concept of "mammal", and "mammal" would be a sub concept of "animal". If the new correspondence (c_j, c_k) is better than the current correspondence (c_j, c_{x_j}), i.e., $\texttt{sim}(c_j, c_k) > \texttt{sim}(c_j, c_{x_j})$, we can replace x_j with k in X. This process enables us to efficiently enhance X while maintaining local consistency. To balance effectiveness and efficiency, we limit the maximum depth of sub concepts in the ontology hierarchical structure to be 3.

Algorithm 2: The newly proposed semantic search.

Input: Source ontology $\mathcal{O}_1$, Target ontology $\mathcal{O}_2$, An individual X
Output: The updated individual X

Set the anchor correspondence set $\Omega = \emptyset$;
for *each concept correspondence* (c_i, c_{x_i}) *where* $(i, x_i) \in X$ **do**
  if $\texttt{sim}(c_i, c_{x_i}) > \rho$ **then**
    $\Omega = \Omega \cup (c_i, c_{x_i})$;
  end
end
for *each anchor correspondence* $(c_i, c_{x_i}) \in \Omega$ **do**
  for *each* c_i*'s sub concept* c_j *in* $\mathcal{O}_1$ **do**
    for *each* c_{x_i}*'s sub concept* c_k *in* $\mathcal{O}_2$ **do**
      if $\texttt{sim}(c_j, c_k) > \texttt{sim}(c_j, c_{x_j})$ **then**
        $x_j = k$;
      end
    end
  end
end
return X;

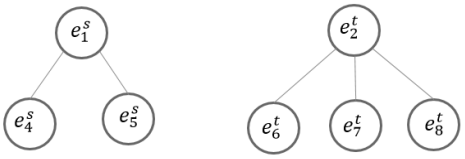
Fig.3 illustrates a representative instance of the semantic search mechanism in operation. Fig.3(a) shows the chromosome of an individual, and Fig.3(b) is the semantic improved solution after the semantic search. Here, we assume that the first gene (e_1^s, e_2^t) is an anchor correspondence. The ontology hierarchical structure is shown in Fig.3(c), from which we can obtain that the sub concepts of e_1^s and e_2^t are $\{e_4^s, e_5^s\}$ and $\{e_6^t, e_7^t, e_8^t\}$, respectively. The similarities of all the possible pairs of the sub concepts are shown in Fig.3(d). Then, we can see that (e_4^s, e_7^t) and (e_5^s, e_8^t) are the best new correspondences. Note that (e_4^s, e_7^t) is already in the current individual, thus is unchanged. However, since the similarity of the the current correspondence $\texttt{sim}(e_5^s, e_6^t) < \texttt{sim}(e_5^s, e_8^t)$, we replace the original correspondence (e_5^s, e_6^t) with (e_5^s, e_8^t) by setting $x_5 = 8$ in the individual. In this way, the semantic search process improves the accuracy of the individual by exploiting the ontology hierarchical structure and updating the correspondences in a way that ensures local consistency.

Gene Value	2	2	10	7	6	3
Gene Index	1	2	3	4	5	6

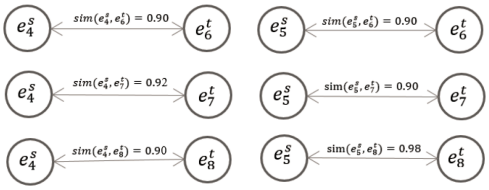
(a) The input individual

Gene Value	2	2	10	7	8	3
Gene Index	1	2	3	4	5	6

(b) The improved individual after the semantic search



(c) Ontology hierarchical structure



(d) Candidate correspondences

FIGURE 3. An example of the semantic search.

The complexity of the semantic search is $O(S^2)$, where $S \ll \min\{|\mathcal{O}_1|, |\mathcal{O}_2|\}$ is the maximal number of (depth-3) sub concepts of the anchor correspondences in the ontologies. It is much more efficient than the traditional local search methods that enumerate the neighborhood, whose complexity is $O(|\mathcal{O}_1| \cdot |\mathcal{O}_2|)$. However, it is still the dominating operator in the algorithm, which is much more time consuming than the selection and genetic operators. For the sake of search efficiency, in each generation of GA-SS, the semantic search is exclusively performed on the population's elite individual.

5. Experimental studies

5.1. **Experiment Design.** The proposed GA-SS is evaluated using the benchmark datasets provided by the Ontology Alignment Evaluation Initiative (OAEI)[1]. These test cases are widely recognized for their coverage of diverse ontology matching scenarios and serve as a standard testbed for assessing the performance of ontology alignment systems. Each test case consists of a pair of ontologies along with a manually curated reference alignment, which allows for quantitative evaluation using established metrics such as precision, recall, and F-measure. It is important to note that these reference alignments are used strictly for evaluation purposes and are not available during the matching process. This setup ensures a fair comparison and realistically reflects the unsupervised nature of matching in practical applications, including engineering domains where manually constructed reference alignments are rarely available.

[1]http://oaei.ontologymatching.org/2016/benchmarks/index.html

The OAEI benchmark dataset is particularly relevant for validating biomedical ontology matching in the context of MCPS. Biomedical ontologies used in MCPS often exhibit diverse forms of heterogeneity, including lexical variations in medical terminology, taxonomic inconsistencies across clinical ontologies, and complex linguistic naming conventions found in electronic health records or anatomical vocabularies. The OAEI benchmarks replicate such challenges, thus providing a rigorous and realistic environment for evaluating the robustness and efficiency of ontology matching methods intended for biomedical integration within MCPS. Table 2 summarizes the benchmark dataset composition, highlighting the main heterogeneity features introduced in different groups of test cases. These cases allow us to systematically assess the adaptability of GA-SS to heterogeneous biomedical ontologies and its potential to support semantic interoperability in MCPS.

TABLE 2. Description of the OAEI Benchmark dataset used for evaluation.

Cases	Brief description
101-104	Two ontologies are nearly identical, serving as a baseline.
201-210	Ontologies differ in lexical and linguistic features.
221-247	Ontologies differ in taxonomic structures.
248-266	Ontologies differ simultaneously in lexical, linguistic, and taxonomic aspects.

The parameter setting of GA-SS is as follows: The population size is set to 20, and the crossover and mutation rates are set to 0.8 and 0.01 according to previous studies in [1]. The maximal number of generations is set to 2000 to ensure the convergence of the algorithm. The new parameters α, θ, ϕ and ρ are set to 0.7, 0.7, 0.7 and 0.55 based on preliminary parameter tuning, which will be shown below.

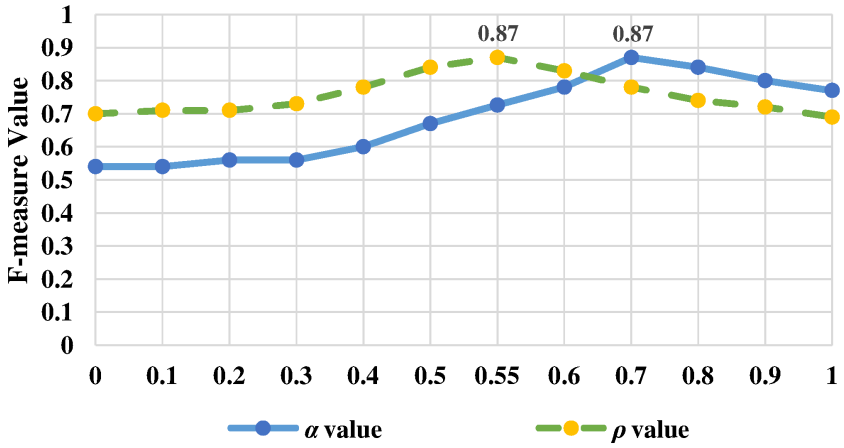


FIGURE 4. The mean F-measure of GA-SS with different α and ρ values over 30 independent runs on all the test cases.

5.2. **Parameter Sensitivity Analysis.** Fig.4 presents the parameter sensitivity analysis, where the horizontal axis denotes parameter values and the vertical axis displays the average F-measure scores across 30 independent GA-SS executions. To isolate parameter effects, all other variables were maintained at their default configurations during individual parameter evaluation. For each parameter under study, the remaining parameters were fixed at their default values to ensure a fair comparison. From Fig.4, we can see that the F-measure is relatively low (below 0.6) for small α values ($\alpha \leq 0.4$). As α increases, the F-measure improves steadily and reaches its peak (0.87) at $\alpha = 0.7$. When α exceeds 0.7, performance gradually declines, dropping to 0.77 when $\alpha = 1.0$. This observation confirms that an appropriate balance between the approximated

F-measure $\mathtt{f}'(\cdot)$ and the locality term yields optimal alignment accuracy. In particular, $\alpha = 0.7$ provides a moderate bias toward $\mathtt{f}'(\cdot)$ while preserving sufficient structural regularization, indicating that GA-SS benefits from combining global alignment accuracy and local consistency constraints.

The parameter ρ defines the similarity threshold for determining whether a correspondence is strong enough to serve as an anchor in the semantic search. When ρ is too small, nearly all correspondences qualify as anchors, effectively degenerating the semantic search into a standard local search. Conversely, when ρ is too large, only highly similar correspondences are retained, limiting exploration. As shown in Fig.4, the optimal performance is achieved when ρ lies between 0.5 and 0.6, with the best result at $\rho = 0.55$. This threshold ensures that the search space is constrained by semantically meaningful correspondences while still allowing sufficient flexibility for discovering new matches.

Overall, the experimental results indicate that GA-SS exhibits low dependence on exact parameter tuning. The optimal settings of $\alpha = 0.7$ and $\rho = 0.55$, derived from the OAEI benchmarks, achieve a stable balance between alignment accuracy, search efficiency, and semantic coherence. These parameters are expected to generalize well across different ontology domains, such as clinical, anatomical, and pharmacological, because the underlying trade-offs they encode, between global and local optimization and between strictness and flexibility in correspondence selection, are inherently domain-independent. This robustness demonstrates the strong generalization capability of GA-SS and its adaptability to heterogeneous biomedical ontology alignment tasks.

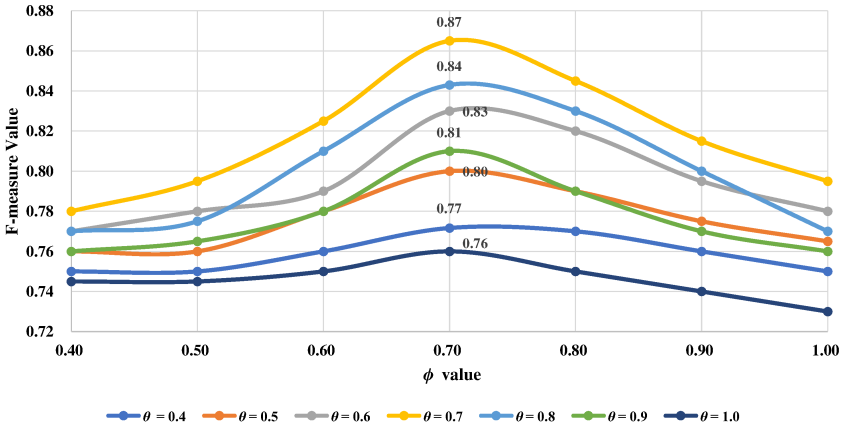


FIGURE 5. The mean F-measure of GA-SS with different θ and ϕ values over 30 independent runs on all the test cases.

Fig.5 shows the relationships between θ and ϕ and their impact on the F-measure value. In particular, the threshold ϕ determines whether the similarity between entities in a specific aspect is significant, while the weight θ quantifies the impact of their similarity difference on the overall score. Together, these two parameters control how GA-SS adaptively balances contributions from syntactic, linguistic, and taxonomic similarities, enabling finer adjustment of the overall similarity estimation. The horizontal axis in Fig.5 indicates the range of ϕ values, whereas the vertical axis displays the corresponding F-measure results. Different colored lines represent different θ values, illustrating how performance varies as ϕ changes. It is evident from the figure that the optimal setting for both θ and ϕ is 0.7, as this configuration yields the highest F-measure of 0.87. Deviation from this configuration in either direction causes a decline in performance, indicating that GA-SS achieves its best balance between flexibility and stability when $\theta = \phi = 0.7$.

This finding also reflects the generalization capability of GA-SS. The optimal combination $\theta = \phi = 0.7$ which provides a moderate weighting among the different similarity dimensions allows

the algorithm to adapt well to diverse ontology structures and terminology styles. Because the OAEI benchmarks encompass various forms of lexical, linguistic, and taxonomic heterogeneity, the same parameter configuration is expected to transfer effectively to other biomedical ontology domains such as clinical, anatomical, or pharmacological ontologies. In these domains, semantic overlap and structural diversity follow similar patterns, meaning that GA-SS can maintain stable performance without domain-specific retuning.

Overall, GA-SS is most sensitive to α (F-measure ranging from 0.54 to 0.87) and least sensitive to ρ (ranging from 0.69 to 0.87). These results confirm the robustness of the framework and its ability to generalize across heterogeneous biomedical domains. In the subsequent experiments, we use $\alpha = \theta = \phi = 0.7$ and $\rho = 0.55$ for GA-SS.

5.3. **Overall Results and Discussions.** The experimental evaluation compares GA-SS against both top-performing OAEI competition participants and current EA-based ontology matching approaches, particularly focusing on the Genetic Algorithm with Stochastic Hill Climbing (GA-SHC) method [43]. The selected comparison methods represent both top-performing OAEI Benchmark solutions and diverse ontology matching techniques, ensuring comprehensive evaluation coverage. The selected OAEI participants for comparative evaluation comprise the following approaches:

- LogMap [50] is a logic-based ontology matching method that adopts a comprehensive evaluation approach based on semantics and vocabulary. It identifies the corresponding ontology mapping relationships through techniques such as ontology modularization and propositional Horn reasoning.
- AOTL [6] implements a greedy algorithmic approach that primarily relies on linguistic-based similarity metrics, which are combined with syntax-based measures through an expert-defined weighted aggregation scheme.
- IAMA [69] as a rule-based matching method, mainly employs syntax-based similarity assessment to identify corresponding entity pairs through query processing.
- YAM++ [33] implements a machine learning-based matching method that integrates multiple similarity metrics through adaptive weight optimization, where the weighting coefficients are determined by either SVM or NB classification models.
- Lily [55] is a machine learning-based matching technique that emphasizes taxonomic similarity measures. It integrates multiple similarity metrics through an adaptive weighting scheme, with coefficients automatically optimized using predefined training data.
- CroMatcher [16] as a rule-based matching approach, performs similarity measure integration by computing the intersection of corresponding alignment sets across multiple similarity categories. The final matching result is extracted using a stable-marriage strategy.

Among them, the performance metrics for comparison methods were acquired directly from the official OAEI benchmark results[2], where the best-performing results of each system were selected from the OAEI 2010-2016 editions corresponding to the same benchmark tasks.

Table 3 presents the comparative F-measure results between GA-SS and current leading ontology alignment approaches on the OAEI Benchmark dataset. The final column displays the statistical summary of GA-SS performance, including mean values and standard deviations calculated across 30 independent experimental trials. To assess statistical significance, we perform Wilcoxon rank sum tests [58] between GA-SS and each baseline algorithm at a significance level of 0.05. The notations "(+)/(−)/(=)" adjacent to competitor results denote whether GA-SS demonstrates statistically superior, inferior, or equivalent performance, respectively, compared to the compared method.

[2]https://oaei.ontologymatching.org

TABLE 3. The comparative analysis presents F-measure results for both GA-SS (including mean and standard deviation across 30 independent runs) and current leading OAEI ontology alignment approaches.

	LogMap	AOTL	IAMA	YAM++	Lily	CroMatch	GA-SS
101	0.95 (+)	0.96 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
201	0.00 (+)	0.94 (+)	0.94 (+)	0.97 (+)	0.99 (+)	1.00 (=)	1.00 (0.00)
202	0.00 (+)	0.84 (+)	0.89 (+)	0.80 (+)	1.00 (-)	0.86 (+)	0.97 (0.01)
221	0.94 (+)	0.95 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
222	0.00 (+)	0.95 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
223	0.94 (+)	0.96 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
224	0.94 (+)	0.95 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
225	0.95 (+)	0.96 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
228	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
232	0.94 (+)	0.96 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
233	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
236	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
237	0.00 (+)	0.95 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
238	0.92 (+)	0.95 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
239	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
240	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
241	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
246	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
247	0.92 (+)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (=)	1.00 (0.00)
248	0.00 (+)	0.30 (+)	0.17 (+)	0.82 (+)	0.89 (+)	0.87 (+)	0.95 (0.01)
249	0.00 (+)	0.30 (+)	0.00 (+)	0.59 (+)	0.73 (+)	0.75 (+)	0.99 (0.01)
250	0.00 (+)	0.32 (+)	0.43 (+)	0.63 (+)	0.84 (+)	0.69 (+)	0.98 (0.03)
251	0.00 (+)	0.51 (+)	0.17 (+)	0.78 (+)	0.86 (+)	0.71 (+)	0.98 (0.01)
252	0.00 (+)	0.31 (+)	0.17 (+)	0.81 (+)	0.91 (+)	0.83 (+)	0.99 (0.01)
253	0.00 (+)	0.31 (+)	0.00 (+)	0.61 (+)	0.78 (+)	0.76 (+)	0.94 (0.01)
254	0.00 (+)	0.29 (+)	0.43 (+)	0.63 (+)	0.60 (+)	0.74 (+)	0.90 (0.02)
257	0.00 (+)	0.33 (+)	0.00 (+)	0.11 (+)	0.22 (+)	0.26 (+)	0.97 (0.02)
258	0.00 (+)	0.50 (+)	0.00 (+)	0.70 (+)	0.62 (+)	0.68 (+)	0.97 (0.02)
259	0.00 (+)	0.31 (+)	0.00 (+)	0.58 (+)	0.72 (+)	0.64 (+)	0.97 (0.01)
260	0.00 (+)	0.50 (+)	0.44 (+)	0.69 (+)	0.58 (+)	0.60 (+)	0.99 (0.02)
261	0.00 (+)	0.31 (+)	0.43 (+)	0.59 (+)	0.26 (+)	0.46 (+)	0.97 (0.01)
262	0.00 (+)	0.29 (+)	0.00 (+)	0.11 (+)	0.16 (+)	0.52 (+)	0.92 (0.00)
+/=/-	32/0/0	24/8/0	15/17/0	15/17/0	14/17/1	14/18/0	

As can be seen from Table 3, GA-SS achieves the highest F-measure across all 32 evaluated test cases, consistently outperforming state-of-the-art OAEI ontology alignment approaches. In the relatively less complex cases (101-247), which involve minor lexical, linguistic, or taxonomic variations, GA-SS attains perfect alignments (F-measure = 1.00) in 18 out of 19 cases. The only exception is case 202, where it still achieves an average F-measure of 0.97, very close to the optimum and ahead of most competitors. These results demonstrate GA-SS's ability to effectively balance exploration and exploitation, ensuring convergence to high-quality solutions even in cases where competing algorithms, whether rule-based, machine learning-based, or logic-based, may become trapped in suboptimal regions of the search space.

In the most challenging subset (248-262), entity names and properties in the target ontology are replaced with random strings, and each entity has significantly fewer instances than in other subsets, severely weakening lexical, structural, and instance-based signals. Although no algorithm achieves optimal alignments under such severe heterogeneity, GA-SS consistently delivers the best performance, with notable margins in cases such as 252, 257, 259, 260 and 261. Across these 13 hardest cases, GA-SS obtains an average F-measure of 0.98, far exceeding the second-best method CroMatcher (0.80) and the third-best Lily (0.78). These results highlight GA-SSs robustness under severe heterogeneity, which is crucial for biomedical MCPS where noisy, incomplete and heterogeneous terminologies are common.

The advantage of GA-SS over the existing matching methods can be attributed to the following factors. First, GA-SS uses three types of similarity measures (syntactic, linguistic, and taxonomic) comprehensively. For example, it significantly outperformed AOTL and IAMA, which use only one or two kinds of similarity measures. Second, GA-SS adaptively adjusts the aggregation pattern of similarity measures based on differences in their computed similarity scores on the entity pair. This makes it much better than the existing methods that treat each similarity measure as a black-box and aggregate them with predefined fixed weights. Thirdly, unlike the compared methods that construct alignments in a single pass, GA-SS employs an iterative improvement process that is further enhanced by the semantic search. This iterative refinement enables GA-SS to continuously search for better alignments, surpassing the alignments produced by the compared methods. Additionally, the semantic search, focusing on refining correspondences by exploring the sub-concepts of each entity, contributes to the algorithm's improved search efficiency. Finally, GA-SS incorporates a local consistency component in the fitness function, which guides the search towards aligning entities that share common neighbors in the ontology. This local consistency property reduces the search space and accelerates convergence towards high-quality solutions.

Table 4 shows the comparison results of GA-SS with GA, which is the counterpart of GA-SS without the proposed semantic search, GA with Stochastic Hill Climbing algorithm (GA-SHC) [10], Compact Multi-tasking Multi-chromosome Genetic Algorithm (CM2GA) [61] and Self-adaptive Designed Genetic Programming (SDGP) [60]. The parameter setting of GA-SHC is consistent with that in its original literature [10], assuming that they have been fine-tuned. The parameters of GA are set the same as GA-SS. For the stopping criteria, in addition to the maximal number of generations (2000), GA-based methods are stopped when the best fitness value is 1.00.

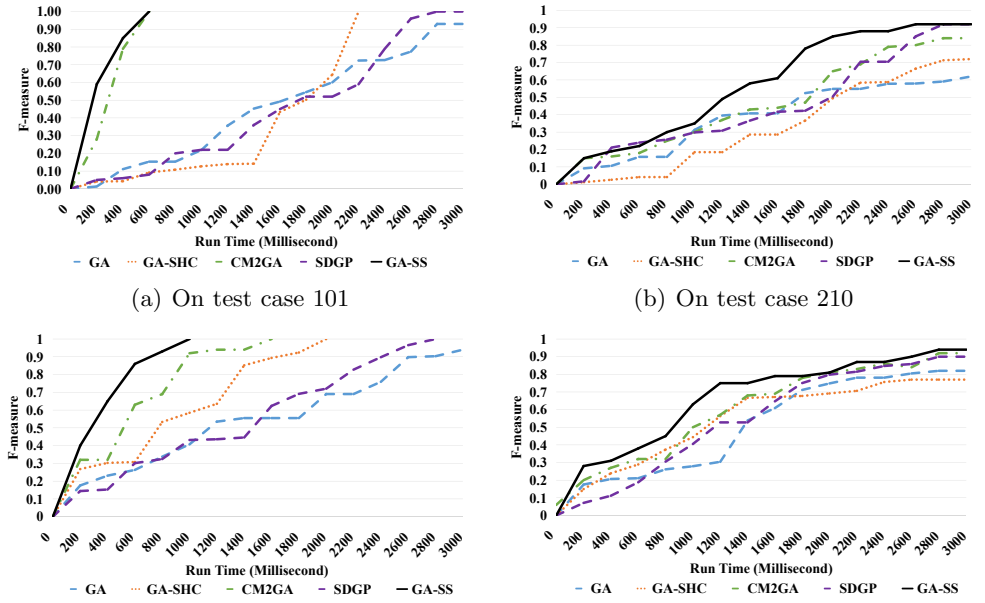
We considered 10 additional test cases 103–104, 203–210 in Table 4, which cannot be included in Table 3 as their results were not available from the OAEI website. For each test case, each compared algorithm is run 30 times independently. The mean and standard deviation of each algorithm is shown in Table 4. Additionally, the Wilcoxon rank-sum test with significance level of 0.05 is used to compare GA-SS with GA and GA-SHC. The "+", "-" and "=" behind the GA and GA-SHC results indicate that GA-SS performed significantly better than, worse than or similar to the corresponding algorithm.

From Table 4, GA-SS consistently demonstrates superior performance across diverse ontology matching tasks with varying degrees of lexical, structural, and linguistic heterogeneity. Compared with its baseline GA (without semantic search), GA-SS achieves significantly better results on all 44 test cases, indicating that the integration of semantic search substantially improves both alignment accuracy and robustness across all dataset types. Against GA-SHC, which augments GA with a stochastic hill climbing procedure, GA-SS records significant gains on 26 cases and achieves statistical parity on the remaining 18, showing that semantic search offers broader and more consistent improvements than local search based refinement, especially on datasets with complex or sparse structural correspondences.

When compared with CM2GA, which employs compact multi-tasking and multi-chromosome encoding to enhance search efficiency, GA-SS attains significant improvements on 16 cases and matches performance on the remaining 28. This indicates that although CM2GA's compact encoding strategy is competitive, semantic search still yields noticeable advantages on tasks with higher semantic ambiguity or weaker lexical cues. Comparing to SDGP, a self-adaptive symbolic regression driven GP framework, GA-SS secures significant improvements on 9 cases, ties on 34, and is marginally worse in only 1 case, demonstrating that semantic search remains complementary to symbolic expression-based similarity construction. In particular, GA-SS maintains high performance even in the most challenging cases (e.g., 248-262) where other methods' results drop sharply, underscoring its robustness under severe heterogeneity.

TABLE 4. The mean (standard deviation) of the F-measure obtained by GA (the counterpart of GA-SS without semantic search), GA-SHC (GA with Stochastic Hill Climbing algorithm), CM2GA (Compact Multi-tasking Multi-chromosome Genetic Algorithm), SDGP (Self-adaptive Designed Genetic Programming) and GA-SS across 30 separate executions on the test cases.

	GA	GA-SHC	CM2GA	SDGP	GA-SS
101	0.93 (0.03) (+)	1.00 (0.01) (=)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00)
103	0.93 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00)
104	0.91 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00)
201	0.85 (0.01) (+)	0.92 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00)
202	0.82 (0.01) (+)	0.85 (0.04) (+)	0.92 (0.01) (+)	0.94 (0.01) (+)	0.97 (0.01)
203	0.91 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00)
204	0.87 (0.03) (+)	0.99 (0.01) (+)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
205	0.88 (0.01) (+)	0.91 (0.02) (+)	0.84 (0.01) (+)	0.96 (0.01) (=)	0.96 (0.01)
206	0.82 (0.02) (+)	0.92 (0.02) (+)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.01)
207	0.85 (0.02) (+)	0.92 (0.02) (+)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.01)
208	0.81 (0.02) (+)	0.92 (0.02) (+)	0.98 (0.01) (+)	0.96 (0.01) (+)	1.00 (0.01)
209	0.64 (0.03) (+)	0.64 (0.03) (=)	0.89 (0.02) (+)	0.90 (0.01) (=)	0.90 (0.01)
210	0.62 (0.02) (+)	0.72 (0.03) (=)	0.84 (0.02) (=)	0.92 (0.01) (=)	0.92 (0.01)
221	0.90 (0.04) (+)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00)
222	0.91 (0.05) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.00) (=)	1.00 (0.00)
223	0.91 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.00)
224	0.89 (0.05) (+)	1.00 (0.00) (=)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.00)
225	0.93 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
228	0.94 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
230	0.90 (0.02) (+)	0.96 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.00)
231	0.90 (0.03) (+)	0.99 (0.01) (+)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
232	0.91 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
233	0.91 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
236	0.89 (0.04) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
237	0.96 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
238	0.85 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
239	0.95 (0.02) (+)	0.98 (0.01) (+)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
240	0.83 (0.06) (+)	0.77 (0.03) (+)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
241	0.81 (0.10) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.01) (=)	1.00 (0.00)
246	0.93 (0.01) (+)	0.98 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.00)
247	0.79 (0.06) (+)	0.76 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.01) (=)	1.00 (0.00)
248	0.82 (0.01) (+)	0.79 (0.03) (+)	0.92 (0.01) (+)	0.85 (0.02) (+)	0.95 (0.01)
249	0.83 (0.04) (+)	0.87 (0.03) (+)	0.99 (0.02) (=)	0.99 (0.02) (=)	0.99 (0.01)
250	0.81 (0.03) (+)	0.79 (0.02) (+)	0.98 (0.01) (=)	0.98 (0.01) (=)	0.98 (0.03)
251	0.86 (0.02) (+)	0.88 (0.02) (+)	0.98 (0.02) (=)	0.86 (0.01) (+)	0.98 (0.01)
252	0.80 (0.04) (+)	0.80 (0.02) (+)	0.99 (0.01) (=)	0.95 (0.02) (+)	0.99 (0.01)
253	0.82 (0.02) (+)	0.77 (0.01) (+)	0.92 (0.02) (+)	0.90 (0.01) (+)	0.94 (0.01)
254	0.86 (0.04) (+)	0.86 (0.03) (+)	0.90 (0.02) (=)	0.90 (0.01) (=)	0.90 (0.02)
257	0.83 (0.03) (+)	0.80 (0.01) (+)	0.97 (0.01) (=)	0.97 (0.01) (=)	0.97 (0.02)
258	0.8 (0.03) (+)	0.78 (0.02) (+)	0.97 (0.02) (=)	0.94 (0.02) (+)	0.97 (0.02)
259	0.82 (0.02) (+)	0.70 (0.02) (+)	0.97 (0.02) (=)	0.97 (0.02) (=)	0.97 (0.01)
260	0.82 (0.05) (+)	0.79 (0.02) (+)	0.97 (0.02) (+)	0.99 (0.02) (=)	0.99 (0.02)
261	0.80 (0.04) (+)	0.81 (0.03) (+)	0.97 (0.02) (=)	0.92 (0.02) (+)	0.97 (0.02)
262	0.85 (0.04) (+)	0.86 (0.02) (+)	0.88 (0.01) (+)	0.92 (0.01) (=)	0.92 (0.00)
+/=/-	44/0/0	26/18/0	16/28/0	9/34/1	



(a) On test case 101 (b) On test case 210

(c) On test case 228 (d) On test case 253

FIGURE 6. The convergence curves of GA (the counterpart of GA-SS without semantic search), GA-SHC (GA with Stochastic Hill Climbing algorithm), CM2GA (Compact Multi-tasking Multi-chromosome Genetic Algorithm), SDGP (Self-adaptive Designed Genetic Programming) and GA-SS on representative test cases.

Fig.6 shows the convergence curves of GA (the counterpart of GA-SS without semantic search), GA-SHC (GA with Stochastic Hill Climbing), CM2GA (Compact Multi-tasking Multi-chromosome Genetic Algorithm), SDGP (Self-adaptive Designed Genetic Programming), and GA-SS on four representative test cases. The x-axis denotes the run time (in milliseconds), and the y-axis represents the F-measure of the best-so-far alignment found by each algorithm. In the simplest case 101, GA-SS and GA-SHC converge rapidly and similarly, both clearly surpassing GA. In case 210, GA-SS consistently outperforms the others, achieving higher accuracy earlier in the run. In case 228, GA-SS reaches the optimal F-measure of 1.0 well before GA-SHC, whereas GA CM2GA, and SDGP fail to match its matching efficiency or accuracy. In the challenging case 253, although none of the methods attain the maximum F-measure, GA-SS achieves notably better convergence rates and final accuracy than the others. Across all cases, GA-SS demonstrates competitive or superior convergence behavior, highlighting its efficiency.

5.4. **Further Analysis.**

5.4.1. *Effectiveness of the New Similarity Measure.* An important new component in GA-SS is the new similarity measure, which adaptively adjusts the weights for aggregating different aspects, i.e., $\texttt{sim}_{sl}$ and $\texttt{sim}_{t}$. To verify the effectiveness of the new similarity measure by Eqns. (6)–(8), we compared it with three commonly used similarity measures: Syntax-based SM (Section 2.2.1), Linguistic-based SM (Section 2.2.2) and Taxonomy-based SM (Section 2.2.3), and aggregation strategies: "Avg", "Max" and "Min". Specifically, in GA-SS, we replace the similarity measure with Syntax-based, Linguistic-based and Taxonomy-based SMs, and the average, maximum, and minimum of $\texttt{sim}_{sl}$ and $\texttt{sim}_{t}$, respectively.

Table 5 shows the mean and standard deviations of the different similarity aggregation strategies in GA-SS over 30 independent runs on the test cases. From the table, we can clearly see that GA-SS significantly outperformed all the other commonly used similarity aggregation strategies on all 44 test cases. This verifies the effectiveness of the newly proposed similarity measure in GA-SS.

TABLE 5. The mean and standard deviation of the F-measure obtained by different similarity aggregation strategies used in GA-SS across 30 separate executions on the test cases.

	Avg.	Max.	Min.	GA-SS
101	0.96 (0.01) (+)	0.84 (0.06) (+)	0.91 (0.03) (+)	1.00 (0.00)
103	0.89 (0.04) (+)	0.89 (0.03) (+)	0.90 (0.04) (+)	1.00 (0.00)
104	0.94 (0.02) (+)	0.89 (0.01) (+)	0.89 (0.05) (+)	1.00 (0.00)
201	0.94 (0.02) (+)	0.92 (0.01) (+)	0.85 (0.05) (+)	1.00 (0.00)
202	0.86 (0.02) (+)	0.86 (0.03) (+)	0.83 (0.04) (+)	0.97 (0.01)
203	0.94 (0.01) (+)	0.87 (0.03) (+)	0.92 (0.04) (+)	1.00 (0.00)
204	0.90 (0.04) (+)	0.89 (0.01) (+)	0.80 (0.02) (+)	1.00 (0.00)
205	0.92 (0.02) (+)	0.91 (0.01) (+)	0.49 (0.03) (+)	0.96 (0.01)
206	0.92 (0.02) (+)	0.78 (0.04) (+)	0.32 (0.08) (+)	1.00 (0.01)
207	0.91 (0.01) (+)	0.82 (0.02) (+)	0.32 (0.07) (+)	1.00 (0.01)
208	0.90 (0.01) (+)	0.73 (0.03) (+)	0.80 (0.02) (+)	1.00 (0.01)
209	0.48 (0.08) (+)	0.47 (0.02) (+)	0.26 (0.10) (+)	0.90 (0.01)
210	0.57 (0.08) (+)	0.47 (0.01) (+)	0.30 (0.09) (+)	0.92 (0.01)
221	0.95 (0.02) (+)	0.81 (0.05) (+)	0.88 (0.03) (+)	1.00 (0.00)
222	0.97 (0.01) (+)	0.84 (0.02) (+)	0.91 (0.02) (+)	1.00 (0.00)
223	0.89 (0.04) (+)	0.84 (0.03) (+)	0.89 (0.04) (+)	1.00 (0.00)
224	0.95 (0.01) (+)	0.85 (0.02) (+)	0.88 (0.04) (+)	1.00 (0.00)
225	0.94 (0.02) (+)	0.86 (0.03) (+)	0.91 (0.02) (+)	1.00 (0.00)
228	0.97 (0.02) (+)	0.81 (0.07) (+)	0.95 (0.01) (+)	1.00 (0.00)
230	0.94 (0.01) (+)	0.76 (0.03) (+)	0.89 (0.03) (+)	1.00 (0.00)
231	0.93 (0.01) (+)	0.85 (0.02) (+)	0.90 (0.04) (+)	1.00 (0.00)
232	0.95 (0.02) (+)	0.86 (0.02) (+)	0.91 (0.03) (+)	1.00 (0.00)
233	0.95 (0.02) (+)	0.83 (0.05) (+)	0.89 (0.04) (+)	1.00 (0.00)
236	1.00 (0.00) (+)	0.88 (0.03) (+)	0.97 (0.02) (+)	1.00 (0.00)
237	0.87 (0.05) (+)	0.80 (0.04) (+)	0.93 (0.02) (+)	1.00 (0.00)
238	0.91 (0.03) (+)	0.79 (0.04) (+)	0.88 (0.04) (+)	1.00 (0.00)
239	0.88 (0.04) (+)	0.84 (0.02) (+)	0.88 (0.05) (+)	1.00 (0.00)
240	0.75 (0.10) (+)	0.65 (0.07) (+)	0.78 (0.06) (+)	1.00 (0.00)
241	0.86 (0.05) (+)	0.86 (0.02) (+)	0.87 (0.03) (+)	1.00 (0.00)
246	0.88 (0.06) (+)	0.81 (0.05) (+)	0.89 (0.04) (+)	1.00 (0.00)
247	0.72 (0.10) (+)	0.77 (0.07) (+)	0.77 (0.05) (+)	1.00 (0.00)
248	0.28 (0.14) (+)	0.30 (0.16) (+)	0.20 (0.20) (+)	0.95 (0.01)
249	0.37 (0.10) (+)	0.41 (0.03) (+)	0.21 (0.15) (+)	0.99 (0.01)
250	0.58 (0.02) (+)	0.43 (0.08) (+)	0.21 (0.21) (+)	0.98 (0.03)
251	0.58 (0.03) (+)	0.57 (0.04) (+)	0.55 (0.11) (+)	0.98 (0.01)
252	0.83 (0.03) (+)	0.80 (0.01) (+)	0.83 (0.03) (+)	0.99 (0.01)
253	0.28 (0.06) (+)	0.29 (0.07) (+)	0.23 (0.03) (+)	0.94 (0.01)
254	0.23 (0.15) (+)	0.23 (0.10) (+)	0.18 (0.20) (+)	0.90 (0.02)
257	0.33 (0.05) (+)	0.34 (0.06) (+)	0.12 (0.03) (+)	0.97 (0.02)
258	0.59 (0.02) (+)	0.59 (0.02) (+)	0.55 (0.02) (+)	0.97 (0.02)
259	0.80 (0.03) (+)	0.82 (0.03) (+)	0.79 (0.03) (+)	0.97 (0.01)
260	0.51 (0.02) (+)	0.45 (0.13) (+)	0.59 (0.09) (+)	0.99 (0.02)
261	0.77 (0.05) (+)	0.77 (0.02) (+)	0.71 (0.08) (+)	0.97 (0.02)
262	0.23 (0.06) (+)	0.23 (0.02) (+)	0.45 (0.00) (+)	0.92 (0.00)
+/=/-	44/0/0	44/0/0	44/0/0	

5.4.2. *Effectiveness of the New Fitness.* The other important new component in GA-SS is the new fitness function. First, it develops an approximated F-measure $\mathtt{f}'(\cdot)$ (Eqns. (2)–(4)) to address the issue of the unavailable true F-measure in practice. Second, it includes an extra term $\mathtt{loc}(\cdot)$ to consider local consistency. To verify the effectiveness of $\mathtt{f}'(\cdot)$ and $\mathtt{loc}(\cdot)$, we compare GA-SS with the following two counterparts:

- GA-SS$_1$: replace the fitness function with $\mathtt{f}'(\cdot)$. Comparing with GA-SS$_1$ can verify the effectiveness of $\mathtt{loc}(\cdot)$;
- GA-SS$_2$: replace the fitness function with $\mathtt{f}(\cdot)+\mathtt{loc}(\cdot)$, where $\mathtt{f}(\cdot)$ is the true F-measure. Comparing with GA-SS$_2$ can verify the effectiveness of $\mathtt{f}'(\cdot)$.

TABLE 6. The mean (standard deviation) of the F-measure derived from GA-SS$_1$, GA-SS$_2$ and GA-SS across 30 separate executions on the test cases.

	GA-SS$_1$	GA-SS$_2$	GA-SS
101	0.95 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
103	0.95 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00)
104	0.94 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
201	0.91 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.00)
202	0.92 (0.01) (+)	1.00 (0.00) (-)	0.97 (0.01)
203	0.96 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00)
204	0.91 (0.04) (+)	1.00 (0.00) (=)	1.00 (0.00)
205	0.90 (0.02) (+)	0.96 (0.01) (=)	0.96 (0.01)
206	0.84 (0.04) (+)	1.00 (0.01) (=)	1.00 (0.01)
207	0.85 (0.02) (+)	1.00 (0.01) (=)	1.00 (0.01)
208	0.88 (0.02) (+)	1.00 (0.03) (=)	1.00 (0.01)
209	0.57 (0.01) (+)	0.90 (0.00) (=)	0.90 (0.01)
210	0.56 (0.01) (+)	0.95 (0.00) (-)	0.92 (0.01)
221	0.91 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.00)
222	0.94 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00)
223	0.90 (0.04) (+)	1.00 (0.00) (=)	1.00 (0.00)
224	0.94 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00)
225	0.91 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
228	0.97 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
230	0.89 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
231	0.92 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
232	0.91 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00)
233	0.85 (0.03) (+)	1.00 (0.00) (=)	1.00 (0.00)
236	0.97 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
237	0.94 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00)
238	0.88 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
239	0.88 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
240	0.91 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
241	0.94 (0.02) (+)	1.00 (0.00) (=)	1.00 (0.00)
246	0.92 (0.01) (+)	1.00 (0.00) (=)	1.00 (0.00)
247	0.82 (0.05) (+)	1.00 (0.00) (=)	1.00 (0.00)
248	0.86 (0.01) (+)	0.97 (0.03) (-)	0.95 (0.01)
249	0.83 (0.02) (+)	0.99 (0.01) (=)	0.99 (0.01)
250	0.92 (0.02) (+)	0.98 (0.02) (=)	0.98 (0.03)
251	0.91 (0.03) (+)	0.98 (0.02) (=)	0.98 (0.01)
252	0.89 (0.02) (+)	0.99 (0.01) (=)	0.99 (0.01)
253	0.89 (0.01) (+)	0.96 (0.01) (-)	0.94 (0.01)
254	0.83 (0.01) (+)	0.90 (0.02) (=)	0.90 (0.02)
257	0.82 (0.02) (+)	0.97 (0.02) (=)	0.97 (0.02)
258	0.87 (0.02) (+)	0.99 (0.02) (-)	0.97 (0.02)
259	0.85 (0.01) (+)	0.97 (0.02) (=)	0.97 (0.01)
260	0.87 (0.04) (+)	0.99 (0.01) (=)	0.99 (0.02)
261	0.82 (0.05) (+)	0.97 (0.02) (=)	0.97 (0.02)
262	0.87 (0.03) (+)	0.96 (0.00) (-)	0.92 (0.00)
+/=/-	44/0/0	0/38/6	

Table 6 shows the comparative results of GA-SS$_1$, GA-SS$_2$ and GA-SS on the test cases. From the table, we can clearly see that GA-SS significantly outperformed GA-SS$_1$ on all 44 test cases. Moreover, without using the locality term, GA-SS$_1$ failed to reach the optimal F-measure on all the test cases. The advantage of GA-SS over GA-SS$_1$ demonstrates the effectiveness of using the locality term in the fitness function. On the other hand, we see that GA-SS showed statistically significantly worse performance than GA-SS$_2$ on 10 out of 44 test cases, while showed no significant difference from GA-SS$_2$ on the remaining 34 test cases. The mean F-measure obtained by GA-SS on all test cases was 0.885, only slightly lower than that of GA-SS$_2$ (0.892). It should be emphasized GA-SS$_2$ uses the true F-measure in the fitness function, which requires the reference alignment to calculate. This makes GA-SS$_2$ not applicable in practice,

where the reference alignment is unknown. The comparison between GA-SS and GA-SS_2 shows the effectiveness of GA-SS and the accuracy of the proposed approximated F-measure $\mathtt{f}'(\cdot)$.

Table 6 compares GA-SS with its two variants, GA-SS_1 and GA-SS_2, across 44 benchmark test cases, revealing clear advantages of the proposed fitness design. GA-SS consistently and significantly outperforms GA-SS_1 on all cases, achieving higher f-measures in both easier datasets (e.g., 101-247, where lexical, linguistic, or taxonomic differences are moderate) and in more challenging ones (e.g., 248-262, characterized by severe heterogeneity and limited lexical/structural cues). Without the locality term $\mathtt{loc}(\cdot)$, GA-SS_1 fails to reach the optimal F-measure in any case, confirming that incorporating local consistency is crucial for capturing valid correspondences and improving alignment quality. Compared to GA-SS_2, which uses the true F-measure in its fitness function and thus has access to reference alignments unavailable in realistic settings, GA-SS matches performance on 38 out of 44 cases and is statistically worse in only 6 cases, mostly within highly degraded datasets where exact F-measure guidance provides a slight edge. In addition, the mean F-measure of GA-SS (0.984) is only marginally lower than GA-SS_2 (0.987), demonstrating that the proposed approximation $\mathtt{f}'(\cdot)$ provides an accurate and practical substitute for the true F-measure, enabling GA-SS to achieve competitive results without requiring inaccessible ground truth.

5.4.3. *Practical Biomedical Ontology Matching.* To further demonstrate the practical utility of GA-SS in biomedical applications, we evaluate it on the OAEI Anatomy track[3], which is widely regarded as a standard benchmark for biomedical ontology alignment. The Anatomy dataset consists of two large-scale reference ontologies: The Adult Mouse Anatomy (MA, 2,744 classes) and the Adult Human Anatomy (HA, 3,304 classes), together with a manually curated gold-standard alignment that enables objective evaluation. We compare GA-SS against six top-performing biomedical ontology matching systems from OAEI, i.e., AML [12], ALIN, Lily, Wikitionary [41], LogMapBio [50] and Matcha [13], as well as two state-of-the-art EA-based techniques, i.e., Geometric Semantic Genetic Programming (GSGP) [63] and Genetic Programming (GP) [49]. GA-SS is configured with the same parameter settings as in the Benchmark dataset experiments, while all competitor methods are run under their original configurations reported in their literature.

Fig.7 presents the comparative results on the Anatomy dataset. GA-SS achieves the highest F-measure (0.98), outperforming both OAEI participants and EA-based methods. Among the OAEI systems, AML (0.86) and ALIN (0.85) perform poorly due to their reliance on predefined lexical and structural heuristics, which cannot capture deeper semantic relations. Lily improves accuracy to 0.90 by incorporating taxonomic similarity, but remains well below GA-SS. Matcha achieves 0.95 through the combination of multiple similarity measures; however, its fixed-weight aggregation limits adaptability, whereas GA-SS integrates these measures dynamically. LogMapBio reaches 0.89 using a logic-based strategy, but at the cost of much higher runtime, whereas GA-SS achieves higher accuracy with a more balanced efficiency. Compared with EA-based competitors, GA-SS also shows clear advantages. GSGP and GP both achieve 0.96, but their runtimes are substantially longer (1,300s and 900s, respectively) than GA-SS (410s). The integration of semantic search and local consistency enables GA-SS to converge to high-quality alignments more efficiently, avoiding the premature convergence that often limits GP-based methods.

Overall, GA-SS achieves a favorable balance between accuracy and efficiency. Although slower than lightweight heuristics such as AML (36s) and Matcha (42s), it is significantly faster than LogMapBio (1,346s) and competitive with EA-based methods, while delivering superior alignment quality. These results suggest that GA-SS can efficiently achieve accurate large-scale biomedical ontology alignment, underscoring its potential for enhancing semantic interoperability in MCPS.

[3]http://oaei.ontologymatching.org/2025/anatomy/index.html

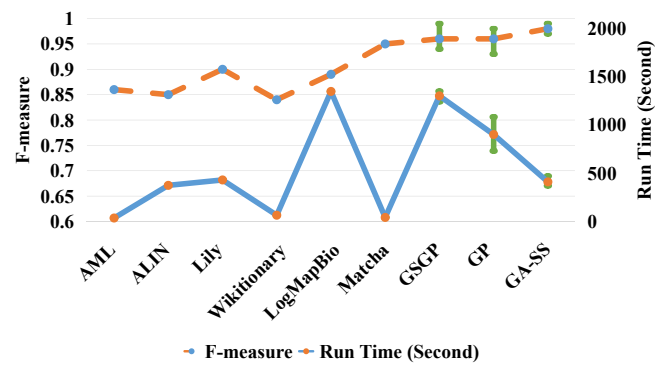


FIGURE 7. The comparative analysis presents F-measure results for GA-SS and state-of-the-art biomedical ontology matching approaches on Anatomy dataset.

6. Conclusions and future work

This paper developed a new GA-SS method to efficiently find high-quality ontology alignments. Compared with traditional GA-based matching approaches, the proposed method has a new fitness function, a novel semantic search strategy and a new hybrid similarity measure. Experimental results confirm the effectiveness of GA-SS with the developed novel fitness function, semantic search mechanism, and similarity measures in enhancing ontology alignment quality. More precisely, in the novel fitness function, the approximated F-measure component offers an accurate estimation of the F-measure, and the local consistent component can better guide the search direction. The semantic search exploits the deep yet well-structured part-whole structures characteristic of ontologies, yielding a substantial performance enhancement over a GA variant that omits SS. The adaptive similarity measure dynamically aggregates lexical, linguistic and taxonomic evidence according to their discriminative power for each ontology pair, which adapts well to various heterogeneous scenarios. Extensive experimental evaluation on OAEI datasets demonstrated that GA-SS consistently outperforms state-of-the-art ontology alignment systems and existing EA-based approaches. In particular, GA-SS achieved near-perfect alignments in relatively simple scenarios and maintained strong robustness in highly heterogeneous cases where competing methods degraded significantly. The results confirm that semantic search and adaptive similarity aggregation jointly enhance both alignment accuracy and computational efficiency, making GA-SS well-suited for real-time and resource-constrained MCPS applications.

Future research will focus on extending GA-SS to support more expressive correspondence types, which are often required in biomedical applications to capture complex relationships like part-whole and functional dependencies. To further improve robustness, we plan to integrate supplementary knowledge sources, including textual descriptions, clinical guidelines, and domain-specific rules, thereby enhancing alignment accuracy when lexical or structural signals alone are insufficient. Another direction is to enhance the semantic search mechanism so that it can dynamically adapt to ontologies with varying structural depth and irregularity, ensuring effective exploration even under highly heterogeneous biomedical scenarios. Finally, we aim to develop adaptive similarity selection strategies that can automatically adjust the relative contributions of syntactic, linguistic, and taxonomic similarity dimensions according to the characteristics of each ontology pair. These advancements will further strengthen the scalability, generalizability, and applicability of GA-SS, enabling reliable semantic interoperability in MCPS for tasks such as electronic health record integration, patient monitoring, and real-time clinical decision support.

7. Acknowledgments

This work was supported by the National Natural Science Foundation of China (No. 62172095), and the authors extend their appreciation to Ongoing Research Funding Program, (ORF-2026-968), King Saud University, Riyadh, Saudi Arabia.

References

[1] Acampora, G., Loia, V., Salerno, S., Vitiello, A. A hybrid evolutionary approach for solving the ontology alignment problem, *Int. J. Intell. Syst.*, Vol.27, No.3, 2012, pp.189-216.

[2] Acampora, G., Loia, V., Vitiello, A. Enhancing ontology alignment through a memetic aggregation of similarity measures, *Inf. Sci.*, Vol.250, 2013, pp.1-20.

[3] Adel, E., El-sappagh, S., Elmogy, M., Barakat, S., Kwak, K.-S. A fuzzy ontological infrastructure for semantic interoperability in distributed electronic health record, *Intell. Autom. Soft Comput.*, Vol.26, No.2, 2020, pp.237-251.

[4] Alves, A., Revoredo, K., Baião, F. A. Ontology alignment based on instances using hybrid genetic algorithm, *In Proc. 7th Int. Conf. Ontol. Match.,* Vol.946, 2012, pp.242243.

[5] Annane, A., Bellahsene, Z., Azouaou, F., Jonquet, C. Building an effective and efficient background knowledge resource to enhance ontology matching, *J. Web Semant.*, Vol.51, 2018, pp.51-68.

[6] Belmehdi, C.B.B., Khiat, A., Keskes, N. OPTIMA: Framework selecting optimal virtual model to query large heterogeneous data, In: Wrembel, R., Gamper, J., Kotsis, G., Tjoa, A.M., Khalil, I. (eds) *Big Data Analytics and Knowledge Discovery,* Lecture Notes in Computer Science, Springer, Cham., Vol.13428, 2022, pp.209-215.

[7] Belhadi, H., Akli-Astouati, K., Djenouri, Y., Lin, J.C.-W. Data mining-based approach for ontology matching problem, *Appl. Intell.*, Vol.50, No.4, 2020, pp.1204-1221.

[8] Berners-Lee, T., Hendler, J., Lassila, O. The semantic web, *Sci. Am.*, Vol.284, No.5, 2001, pp.34-43.

[9] Biniz, M., Fakir, M. An ontology alignment hybrid method based on decision rules, *Int. Arab J. Inf. Technol.*, Vol.16, No.6, 2019, pp.1114-1120.

[10] Chu, S.-C., Xue, X.S.,Pan, E.Y., Wu, X.J. Optimizing ontology alignment in vector space, *J. Internet Technol.*, Vol.21, No.1, 2020, pp.15-22.

[11] Euzenat, J., Meilicke, C., Stuckenschmidt, H., Shvaiko, P., Trojahn, C. Ontology alignment evaluation initiative: Six years of experience In: Spaccapietra, S. (eds) *Journal on Data Semantics XV,* Lecture Notes in Computer Science, Vol 6720, Springer, Berlin, Heidelberg, Vol.15, 2011, pp.158-192.

[12] Faria, D., Santos, E., Balasubramani, B.S., Silva, M.C., et al. AgreementMakerLight, *Semant. Web*, Vol.16, No.2, 2025, pp.SW–233304.

[13] Faria, D., Silva, M.C., Cotovio, P., Ferraz, L., et al. Results in OAEI 2024 for Matcha, In: *CEUR Workshop Proc.*, Vol.3897, CEUR-WS, 2024, pp.110-117.

[14] Ferranti, N., Soares, S.S.R.F., de Souza, J.F. Metaheuristics-based ontology meta-matching approaches, *Expert Syst. Appl.,* Vol.173, 2021, p.114578.

[15] Fürst, J., Argerich, F. Mauricio, Bin, Ch. VersaMatch: Ontology matching with weak supervision, In: *Proc. VLDB Endow.*, Vol.16, No.6, 2023, pp.1305-1318,

[16] Gulic, M., Vrdoljak, B., Ptivcek, M. CroMatcher 2.0: A comprehensive analysis of the improved ontology matching system, *IEEE Access*, Vol.13, 2025, pp.27610-27640.

[17] Horton, J.D., Kilakos, K. Minimum edge dominating sets, *SIAM J. Discrete Math.*, Vol.6, No.3, 1993, pp.375-387.

[18] Hossain, M.M., Kashem, M.A., Nayan, N.M., Chowdhury, M.A. A Medical Cyber-physical system for predicting maternal health in developing countries using machine learning, *Healthc. Anal.*, Vol.5, 2024, pp.100285.

[19] Houshmand, M., Naghibzadeh, M., Araban, S. Reliability-based similarity aggregation in ontology matching, In: *Proc. IEEE Int. Conf. Intell. Comput. Intell. Syst.*, Vol.3, 2010, pp.744-749.

[20] Ji, Q., Haase, P., Qi, G. Combination of similarity measures in ontology matching using the OWA operator. In: Yager, R.R., Kacprzyk, J., Beliakov, G. (eds) *Recent Developments in the Ordered Weighted Averaging Operators: Theory and Practice. Studies in Fuzziness and Soft Computing,* Springer, Berlin, Heidelberg, Vol.265, 2011, pp.281-295.

[21] Ji, Q., Liu, W., Qi, G., Bell, D.A. LCS: A linguistic combination system for ontology matching, In: Lang, J., Lin, F., Wang, J. (eds) *Knowledge Science, Engineering and Management,* Lecture Notes in Computer Science, Springer, Berlin, Heidelberg, Vol. 4092. 2006, pp.176-189.

[22] Khater, H.M., Sallabi, F., Serhani, M.A., Barka, E., et al. Empowering healthcare with cyber-physical system a systematic literature review, *IEEE Access*, Vol.12, 2024, pp.83952-83993.
[23] Khoudja, M.A., Fareh, M., Bouarfa, H. Deep embedding learning with auto-encoder for large-scale ontology matching, *Int. J. Semant. Web Inf. Syst.*, Vol.18, No.1, 2022, pp.1-18.
[24] Laadhar, A., Ghozzi, F., Megdiche, I., Ravat, F., et al. POMap++ results for OAEI 2019: Fully automated machine learning approach for ontology matching, In: *Proc. 14th Int. Workshop Ontol. Matching*, 2019, pp.169-174.
[25] Lin, F.Y., Sandkuhl, K. A survey of exploiting WordNet in ontology matching, In: *Artif. Intell. Theory Pract. II, IFIP 20th World Comput. Congr.*, Springer, 2008, pp.341-350.
[26] Ling, H.F., Wu, J., Li, P., Shen, J.L. Attention-aware network with latent semantic analysis for clothing invariant gait recognition, *Comput. Mater. Continua*, Vol.60, No.3, 2019, pp.1041-1054.
[27] Mahdipour, E., Ghasemzadeh, M. Biological network alignment using hybrid genetic algorithm and simulated annealing, In: *Proc. 10th Int. Conf. Comput. Knowl. Eng. (ICCKE)*, IEEE, 2020, pp.215-220.
[28] Malik, S., Jain, S. Deep convolutional neural network for knowledge-infused text classification, *New Gener. Comput.*, Vol.42, No.1, 2024, pp.157-176.
[29] Mani, S., Annadurai, S. Explicit link discovery scheme optimized with ontology mapping using improved machine learning approach, *Stud. Inform. Control*, Vol.30, No.1, 2021, pp.67–75.
[30] Miller, G.A. WordNet: A lexical database for English, *Commun. ACM*, Vol.38, No.11, 1995, pp.39-41.
[31] Mohammadi, M., Hofman, W., Tan, Y.-H. Simulated annealing-based ontology matching, *ACM Trans. Manage. Inf. Syst.*, Vol.10, No.1, 2019, pp.1-24.
[32] Moroz, O., Stepashko, V. New two-parametric mutation operator for inductive modelling using combinatorial-genetic algorithm, *Proc. 12th Int. Conf. Adv. Comput. Inf. Technol. (ACIT)*, IEEE, 2022, pp.76-79.
[33] Ngo, D.-H., Koopman, B. Self-learning ontological concept representation for searching and matching tasks, In: *Proc. Int. Semant. Web Conf. (ISWC)*, CEUR-WS, 2022, pp.73-78.
[34] Nguyen, T.T.A., Conrad, S. Ontology matching using multiple similarity measures, In: *Proc. 7th Int. Joint Conf. Knowl. Discov. Knowl. Eng. Knowl. Manage. (IC3K)*, IEEE, 2015, pp.603-611.
[35] Nkisi-Orji, I., Wiratunga, N., Massie, S., Hui, K.-Y., Heaven, R. Ontology alignment based on word embedding and random forest classification, In: Berlingerio, M., Bonchi, F., Gärtner, T., Hurley, N., Ifrim, G. (eds) *Machine Learning and Knowledge Discovery in Databases,* ECML PKDD 2018, Lecture Notes in Computer Science, Springer, Vol.11051, 2019, pp.557-572.
[36] Okabe, M., Tsuchida, J., Yadohisa, H. F-measure maximizing logistic regression, *Commun. Stat. Simul. Comput.*, Vol.53, No.5, 2024, pp.2554-2564.
[37] Ouali, I., Ghozzi, F., Taktak, R., Sassi, M.S.H. Ontology alignment using stable matching, *Procedia Comput. Sci.*, Vol.159, 2019, pp.746-755.
[38] Pardeshi, S., Billakurthi, S.K., Gawade, S. Comprehensive review of population based metaheuristic algorithms & ontology integration, In: *Proc. 2nd Int. Conf. Emerg. Trends Inf. Technol. Eng. (ICETITE)*, IEEE, Vellore, India, 2024, pp.1-8.
[39] Peng, W., Zhou, Y., Xu, B. Matching large ontologies based on reduction anchors, In: *Proc. 22nd Int. Joint Conf. Artif. Intell. (IJCAI)*, 2011, pp.2344-2348.
[40] Pezoulas, V.C., Kalatzis, F., Exarchos, T.P., Goules, A., et al. Dealing with open issues and unmet needs in healthcare through ontology matching and federated learning, *In: Jarm, T., Cvetkoska, A., Mahnič-Kalamiza, S., Miklavcic, D. 8th Eur. Med. Biolog. Eng. Conf.,* IFMBE Proc., Springer, Cham., Vol.80., 2020, pp.306-313.
[41] Portisch, J., Hladik, M., Paulheim, H. Wiktionary matcher, In: *CEUR Workshop Proc.*, RWTH Aachen, Vol.2536, 2020, pp.181-188.
[42] Qazvinian, V., Abolhassani, H., Haeri, S.H., Hariri, B.B. Evolutionary coincidence-based ontology mapping extraction, *Expert Syst.*, Vol.25, No.3, 2008, pp.221-236.
[43] Ryma, G., Mohamed-Khireddine, K. Genetic algorithm with hill climbing for correspondences discovery in ontology mapping, *J. Inf. Technol. Res.*, Vol.12, No.4, 2019, pp.153-170.
[44] Sankoff, D. *Time Warps, String Edits, and Macromolecules: The Theory and Practice of Sequence Comparison,* Center for the Study of Language and Inf, 1999, 408 p.
[45] Sentürk, F., Aytac, V. A graph-based ontology matching framework, *New Gener. Comput.*, Vol.42, No.1, 2023, pp.33-51.
[46] Silva, M.C., Faria, D., Pesquita, C. Matching multiple ontologies to build a knowledge graph for personalized medicine, In: Groth, P., et al. *The Semantic Web., Lecture Notes in Computer Science,* Springer, Cham., Vol.13261, 2022, pp.461-477.
[47] Solimando, A., Jimenez-Ruiz, E., Guerrini, G. Minimizing conservativity violations in ontology alignments: Algorithms and evaluation, *Knowl. Inf. Syst.*, Vol.51, No.3, 2017, pp.775-819.
[48] Stoilos, G., Stamou, G., Kollias, S. A string metric for ontology alignment, In: *Int. Semant. Web Conf. (ISWC)*,Springer, 2005, pp.624-637.

[49] Sun, D., Lv, Q., Tsai, P.-W., Xue, X., Zhang, K. Anchor-based ontology partitioning and Genetic Programming with Relevance Reasoning for large-scale biomedical ontology matching, *Expert Syst. Appl.*, Vol.270, 2025, pp.126445.

[50] Teymurova, S., Jimenez-Ruiz, E., Weyde, T., Chen, J. OWL2Vec4OA: Tailoring knowledge graph embeddings for ontology alignment, In: *Int. Knowl. Graph Semant. Web Conf.*, Springer, Vol.15459, 2024, pp.168-182.

[51] Touati, C., Kemmar, A. Deep reinforcement learning approach for ontology matching problem, *Int. J. Data Sci. Anal.*, Vol.18, No.1, 2024, pp.97-112.

[52] Wang, J., Cheng, Z., Ersoy, O.K., Zhang, P., Dai, W. Multi-offspring genetic algorithm with two-point crossover and the relationship between number of offsprings and computational speed, *J. Comput.*, Vol.30, No.5, 2019, pp.111-127.

[53] Wang, J., Ding, Z., Jiang, C. GAOM: Genetic algorithm-based ontology matching, *IEEE Asia-Pacific Conference on Services Computing (APSCC'06),* Guangzhou, China, 2006, pp.617-620.

[54] Wang, N.-N., Zhu, B., Li, X.-L., Liu, S., et al. Comprehensive review of drug–drug interaction prediction based on machine learning: Current status, challenges, and opportunities, *J. Chem. Inf. Model.*, Vol.64, No.1, 2023, pp.96-109.

[55] Wang, P., Hu, Y. Matching biomedical ontologies via a hybrid graph attention network, *Front. Genet.*, Vol.13, 2022, pp.893409.

[56] Wang, Z. A semi-supervised learning approach for ontology matching, In: *Proc. Chinese Semant. Web and Web Sci. Conf. (CSWC)*, Springer, 2014, pp.17-28.

[57] Winkler, W.E. The state of record linkage and current research problems, In: *Proc. Surv. Methods Sect.*, Springer, 1999, pp.73-90.

[58] Woolson, R.F. Wilcoxon signed-rank test, *Encyclopedia of Biostatistics,* John Wiley & Sons, Ltd, Vol.2007, 2007, pp.1-3.

[59] Wu, X., Xue, X., Hu, W. Argumentation based ontology alignment extraction, In: *Int. Conf. Adv. Mach. Learn. Technol. Appl.*, Springer, Cham., Vol.1339 2021, pp.1028-1037.

[60] Xue, X. Automatic knowledge graph matching via self-adaptive designed genetic programming, *Knowl.-Based Syst.*, Vol.293, 2024, pp.111628.

[61] Xue, X., Lin, J. C.-W., Su, T. Compact multitasking multichromosome genetic algorithm for heuristic selection in ontology matching, *IEEE Trans. Artif. Intell.*, Vol.5, No.12, 2024, pp.6752-6766.

[62] Xue, X., Pan, J.-S. An overview on evolutionary algorithm based ontology matching, *J. Inf. Hiding Multimed. Signal Process.*, Vol.9, No.1, 2018, pp.75-88.

[63] Xue, X., Sun, D., Shankar, A., Viriyasitavat, W., Siarry, P. Efficient large-scale biomedical ontology matching with anchor-based biomedical ontology partitioning and compact geometric semantic genetic programming, *J. Ind. Inf. Integr.*, Vol.41, 2024, pp.100637.

[64] Xue, X., Wang, Y. Optimizing ontology alignments through a memetic algorithm using both matchfmeasure and unanimous improvement ratio, *Artif. Intell.*, Vol.223, 2015, pp.65-81.

[65] Xue, X., Yang, C., Mao, G., Zhu, H. Semi-automatic ontology matching based on interactive compact genetic algorithm, *Int. J. Pattern Recognit. Artif. Intell.*, Vol.36, No.05, 2022, pp.2257002.

[66] Xue, X., Ye, M., Nian, Q. Matching ontologies with kernel principle component analysis and evolutionary algorithm, *Demonstr. Math.*, Vol.57, No.1, 2024, pp.20240021.

[67] Yuan, C., Yang, H., Pan, Y. Research on data link ontology mapping algorithm based on bayesian network model, *IEEE Access*, Vol.7, 2019, pp.185698-185709.

[68] Zhang, F., Guo, Z. A novel optimization method for ontology matching based on heuristic population evolution algorithm, *Arab. J. Sci. Eng.*, Vol.44, No.4, 2019, pp.3137-3153.

[69] Zhang, Y., Wang, X., He, S., Liu, K., et al. IAMA results for OAEI 2013, In: *Proc. Int. Semant. Web Conf. (ISWC)*, CEUR-WS, Vol.1111, 2013, pp.123-130.

[70] Zou, S., Liu, J., Yang, Z., Hu, Y., Wang, P. Lily results for OAEI 2021, In: *Proc. Int. Semant. Web Conf. (ISWC)*, CEUR-WS, 2021, pp.167-174.



Xingsi Xue - is a Professor at the College of Artificial Intelligence, Yango University, and a Researcher at the School of Computer Science and Mathematics, Fujian University of Technology. He received his Ph.D. degree from Xidian University, China, in 2014. He was a postdoc fellow at School of Engineering and Computer Science, Victoria University of Wellington, New Zealand, in 2022 and 2023. His current research interests include knowledge and data engineering, intelligent computation and machine learning. He is a Senior Member of IEEE and a Distinguished Member of China Computer Federation (CCF).



Syed Muhammad Waqas - is an Associate Professor with the College of Artificial Intelligence at YanGo University, China. He received his Ph.D. degree in Computer Science and Technology from Xian Jiaotong University, China, in 2024. His research interests include 5G/6G wireless communications, internet of vehicles, machine learning, artificial intelligence, unmanned aerial vehicles, resource allocation, cloud computing, cybersecurity, and the internet of things. He is a member of IEEE and the China Computer Federation (CCF).



Chien-Ming Chen - is a Professor with the School of Artificial Intelligence (School of Future Technology), Nanjing University of Information Science and Technology, Nanjing, China. He received his Ph.D. degree from National Tsing Hua University, Hsinchu, Taiwan, in 2010. His current research interests include AI, blockchain, network security, big data, and data mining.



Mohammed Amoon - is currently a Professor with the Department of Computer Science, King Saud University. He received his B.Sc. degree in Electronic Engineering and M.Sc. and Ph.D. degrees in Computer Science and Engineering from Menoufia University in 1996, 2001, and 2006, respectively. His research interests include agent-based systems, fault tolerance techniques, scheduling algorithms, grid computing, cloud computing, fog computing, healthcare informatics, and the Internet of Things (IoT).