International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Quantum Machine Learning Technique for Automatic Retrosynthetic Reaction Pathway Search Method

Jagannath Jijaba Kadam[1*], Ganpati Martand Kharmate[2], Mahadeo Ramchandra Jadhav[3], Siddhanath Abasaheb Howal[4], Vikram Uttam Pandit[5]
[1*]Professor, Chemistry Department, Bharati Vidyapeeth college of Engineering, Navi Mumbai, India, Email: jjkadam702@gmail.com
[2]Assistant Professor, Physics Department, Bharati Vidyapeeth college of Engineering, Navi Mumbai, India, Email: ganpati.kharmate@bvcoenm.edu.in
[3]Assistant Professor, Chemistry Department, Bharati Vidyapeeth college of Engineering, Navi Mumbai, India, Email: mahadeojadhav2013@gmail.com
[4]Assistant Professor, Gharada Institute of Technology, A/P Lavel, khed, Ratnagiri, Maharashtra, India, Email: siddhu.howal@gmail.com
[5]Assistant Professor, Haribhai V. Desai College, Pune-411002, Maharashtra, India, Email:vikramupandit@gmail.com

Abstract: Retrosynthetic analysis often involves evaluating many potential candidate reaction pathways and molecules at multiple stages of the reaction, resulting in complex retrosynthesis trees that need to be searched and parsed efficiently. Computational approaches could significantly aid the chemist in solving different aspects of the retrosynthesis problem, such as the graph-theoretic search methodologies for efficient tree traversal to identify feasible reaction pathways, dictionary-based methods to evaluate a large search space of precursors, and chemistry-driven heuristics to eliminate practically infeasible routes. In this research, a new single-step retrosynthesis prediction method of the Retro TRAE SMILES-based translation technique is proposed. Accordingly, quantum computing with tree-tensor network topology is presented to construct an automatic data-driven end-to-end retrosynthetic route planning system (Auto-Syn-Route), which is presented based on the heuristic scoring function. AutoSynRoute successfully reproduced published synthesis routes for the four case products. The model is trained in an end-to-end and fully data-driven fashion. Unlike previous models translating the SMILES strings of reactants and products, a new way of representing a chemical reaction based on molecular fragments is introduced. It is demonstrated that the new approach yields better prediction results than current state-of-the-art computational methods. The new approach resolves the major drawbacks of existing retrosynthetic methods such as generating invalid SMILES strings. The proposed method is implemented using Python software. The proposed approach predicts highly similar reactant molecules with an accuracy of 68%. In addition, the proposed method yields more robust predictions than existing methods. However, the experiments demonstrate that the proposed scheme significantly improves the success rate of solving the retrosynthetic problem by 97% while maintaining the performance of the quantum tree tensor for predicting valid reactions.

Keywords: Retrosynthetic, RetroTRAE, SMILES, Single-Step Retrosynthesis Prediction, Quantum Computing, Tree-Tensor Network.

I. INTRODUCTION

Retrosynthetic planning is a fundamental problem in chemistry for finding a pathway of reactions to synthesize a target molecule. Well-planned and practical retrosynthetic pathways are essential for the effective and ecologically sustainable synthesis of important compounds. Robert Robinson introduced retrosynthetic analysis in the tropinone synthesis process and E. To create target compounds, organic chemists employ a fundamental method called J. Corey. A molecule's production process is typically varied, especially for complex substances like natural products. Based on a collection of reaction rules, choices for each transformation are developed, and a variety of optimization algorithms then suggest potential reaction paths. Even if computer-assisted retrosynthetic route planning and reaction prediction have made significant strides, completely data-driven autonomous retrosynthetic route planning is still difficult [1] [2]. Retrosynthesis is likely one of the more challenging processes among the several activities involved. Retrosynthesis involves designing effective synthetic routes for a certain target. The necessity to identify a series of disconnections schemes, appropriate building blocks, and effective group protection techniques are some of the main justifications.

For a long time, the most effective method used in computer programmes was rule-based or similarity-based. These approaches do not learn chemistry from data, but rather codify synthon creation rules, even though they indicate relatively efficient pathways to molecules of interest. Rule-based systems' fundamental flaw is the requirement for time-consuming manual encoding, which prevents growth as data set sizes grow. Additionally, when more rules are codified, it becomes more

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

difficult to determine whether all of the current rules and the new ones are logically consistent with one another. Eventually, this complexity could lead to an insoluble dilemma [2]. Retrosynthesis assisted by artificial intelligence (AI) aims to automate this procedure by inferring new predictions from past chemical reactions. Even though several models have shown that they can be used for automated retrosynthesis, much more work needs to be done to improve prediction accuracy to a level that is more useful [3]. A molecule's production process is typically varied, especially for complex substances like natural products. Planning a target molecule's efficient and environmentally friendly path in some way heavily depends on the expertise of professional chemists [4].

Nevertheless, apart from the use of some straightforward heuristics, it is difficult to systematically rank strategic pathways due to the lack of a comprehensive pathway evaluation mechanism. A dynamic tree-structured long short-term memory (tree-LSTM) model is utilized in [5] to assess the relative strategic levels of retrosynthesis pathways. A unique Graph Transformer architecture is created to adaptively learn discriminative and chemically relevant molecule representations by incorporating chemical knowledge as prior information, showcasing the great capability in molecule feature representation learning [6]. However, these are currently too sluggish to be employed for virtual screening procedures that identify possible bioactivity before screening the synthetic viability of millions of created or enumerated molecules. Here, we present a method based on machine learning (ML) that can determine whether such a synthetic route may be determined for a given molecule [7]. In advance, depth-first proof-number search (DFPN) algorithm [8], deep highway networks [9] and deep learning models [10] are investigated the retrosynthesis model. With off-policy data, it develops a neural search bias while maintaining the search as an AND-OR tree. Then, under the guidance of this neural network, it efficiently executes the best-first search during fresh planning episodes [11].

Machine learning methods frequently concentrate on presenting high-accuracy statistics for the one-to-one mapping of molecules in reaction data to the template extracted from the observed event [12] when prioritising reaction templates or molecular transformations. Fully data-driven automatic retrosynthetic route design is still difficult, despite significant advancements in computer-assisted retrosynthetic route planning and reaction prediction. With the help of the multi-head attention-based transformer architecture, that has proven effective in machine translation tasks and each reaction prediction challenge as a data-driven sequence-to-sequence problem [13]. Substructure-level decoding models are proposed [14] in which the substructures can be automatically extracted using a fully data-driven technique and are reaction-aware. The performance can be increased much more if the substructure extraction's accuracy is increased. Despite the existence of reaction databases, it is challenging to explore reaction information, leading to a path explosion issue because of the vast search area. Based on a hybrid generative exploration and exploitation of reaction knowledge graphs storing massive data of patented reactions ML-based retrosynthetic prediction, an AI system that facilitates synthetic path creation at the fundamental stages of research and process design [15]. The present work reduces the gap between the widely used supervised learning of single-step retrosynthetic models and the goal of retrosynthetic planning. The proposed work presented a Retro TRAE SMILES-based translation technique for a single-step retrosynthesis prediction method. However, quantum computing with tree-tensor network topology is used for constructing an automatic data-driven end-to-end retrosynthetic route planning system. The rest of the papers are organized as follows, section 2 reveals the literature survey of the study, and section 3 exhibits the problem definition and motivation of the study. Section 4 portrays the proposed research methodology, section 5 displays the experimentation and result discussion, and section 6 demonstrated the research conclusion.

II. Literature Survey

An ongoing difficulty in organic synthesis is creating effective synthetic routes for a given target molecule. Atom environments are the perfect, independent, chemically significant building pieces that offer a high-resolution molecular depiction. In contrast to other cutting-edge neural machine translation-based methods, Ucak *et al* [16] has been developed a new single-step retrosynthesis concept called RetroTRAE, which was free from all SMILES-based translation problems, achieves top-1 accuracy of 58.3% on the USPTO test dataset and top-1 accuracy of 61.6% when highly similar analogues are taken into account. To employ fragmental and topological descriptors as natural inputs for retrosynthetic prediction tasks, Their methodology proposed a novel scheme. The choice and creation of appropriate synthetic pathways are crucial decisions that influence the productivity and economics of chemical operations, including reactions and the search for novel compounds. Even with reaction databases, exploring response information is challenging, leading to path explosion because of the vast search space and opposing constraints like economics, safety, efficiency, etc. As a result, Jeong *et al* [17] proposed the ASICS (Advanced System for Intelligent Chemical Synthesis) intelligent system, which supports synthetic path design based on the hybrid generative exploration and exploitation base, and ML-based retrosynthetic prediction.

Computer-aided retrosynthesis has the potential to help chemists build synthetic routes, although at the moment it is laborious and yields findings of poor quality. To carry out a retrosynthesis prediction task taught by using the Transformer neural network architecture, Zheng *et al* [18] constructed a template-free self-corrected retrosynthesis predictor. The method converts the planning of retrosynthesis into a machine translation issue between the reactants' and products' molecular linear notations. The high-throughput synthesis technique is hampered by the cost and time involved in this type of technology. Therefore, a framework for retrosynthetic analysis was developed in [19] using hybrid reaction templates and GC-based thermodynamic models. Using the Breadth-First Search (BFS) algorithm, putative retrosynthesis paths that were thermodynamically viable were found. To demonstrate the viability and dependability of the suggested framework, three case studies using aspirin, ibuprofen, and zatosetron were presented.

The RetroPrime approach, proposed forth by Wang *et al* [20], integrates the chemists' retrosynthetic strategy of breaking down a molecule into synthons and producing reactants by adding leaving groups. For these two levels, flexible Transformer models were used. It was previously recognised

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

that the Transformer-based retrosynthesis model's outputs frequently suffer from a lack of diversity and a high degree of chemical implausibility. RetroPrime was made to address these issues. The suggested framework can be used to determine synthesis routes that take thermodynamic feasibility into account. To demonstrate the viability and dependability of the suggested framework [21], two case studies utilising aspirin and ibuprofen were presented.

Zheng *et al* [22] created BioNavi-NP, a navigable and user-friendly toolset, to forecast the biosynthetic routes for both NPs and NP-like chemicals. First, end-to-end transformer neural networks are used to train a single-step bio-retrosynthesis prediction model employing both conventional organic and biosynthetic reactions. The toolbox, curated datasets, and learning models were all freely available to help with the reconstruction and elucidation of NPs' biochemical processes. The deterministic inference, which contradicts the concept that many compounds can be produced through a variety of reaction types with distinct sets of reactants, causes the majority of them to have difficulty identifying various chemical reactions for the desired outcome. He *et al* [23] used the discrete latent variables to boost reaction diversity and produce a range of reactants. This led to the development of a novel sequence-based method called RetroDVCAE, which integrates conditional variational autoencoders into single-step retrosynthesis and links discrete latent variables to the generation procedure.

A data-driven CASP application created by Ishida *et al* [24] incorporates numerous retrosynthesis knowledge components and presented the information as programmable factors in the assessment of potential search routes. According to the experimental findings, ReTReK successfully sought synthetic routes depending on the supplied retrosynthesis knowledge, showing that these routes were preferred over those that were not. It was anticipated that a data-driven CASP application will improve the performance of both completed and upcoming data-driven CASP applications. Jiang *et al* [25] offered a chance to address some common problems, such as the need for substantial expertise, the sub-optimality of routes, and the high cost of calculation. Also, they have discussed the state of AI-driven retrosynthesis prediction at the moment [26, 28]. Then go through related AI methods and new developments that allow for retrosynthesis prediction [29-30].

III. Research Problem Definition And Motivation

Planning the reaction pathways of organic molecules is a central component of organic synthesis. The idea of reducing the complexity of a desired organic molecule by considering all logical disconnections forms the basis of the retrosynthetic approach [31-34]. The retrosynthetic approach aims to suggest a logical synthetic route to generate a target molecule from a set of available reaction building blocks [35-37]. A conventional retrosynthetic approach acts recursively on a target molecule until chemically reasonable pathways are identified. From a broader perspective, existing predictors for forward and backward reactions can be classified into those that rely on known reaction templates and those that are template-free, data-driven networks trained in an end-to-end fashion [38-39].

Template-free methods have emerged as an effective means to complement the following issues of template-based methods. Exploring the space of possible reaction templates is challenging because of the vast size of the chemical space [40-42]. If only a limited number of reaction templates are used, template-based methods may not be able to provide novel disconnections [43-45]. On the contrary, if a large number of reaction templates are considered, the computational burden to find a proper template increases significantly [46-47]. Currently, templates are either hand-crafted by experts or generated from reaction databases with heuristic algorithms. Thus, the degree of template generality/specificity can lead to either low-quality or incomplete recommendations [48-49]. Lastly, reaction templates are extracted based on atom mapping, which remains a challenging issue for all template-based methods. Atom mapping quality also affects model performance. Considering the complexity of retrosynthetic analysis, an efficient representation of source-target data structure is critical for accurate predictions. In this study, the research shows that representing molecules using sets of atom environments (AE) is an efficient alternative approach to conventional SMILES-based approaches for devising a retrosynthetic prediction model.

IV. Proposed Research Methodology

Automation of the task of obtaining transition state structures and related reaction paths has been one of the significant subjects in computational chemistry. Recent advances made it possible to construct a complex reaction path network consisting of thousands or more paths based on DFT calculations. Consequently, the study mimics chemical reasoning and predicts reactant candidates by learning the changes in atom environments associated with the chemical reaction. Through careful inspection of reactant candidates, the study demonstrates atom environments as promising descriptors for studying reaction route prediction and discovery. Here, the research proposed a new single-step retrosynthesis prediction method, of Retro TRAE SMILES-based translation technique. Further, constructed an automatic data-driven end-to-end retrosynthetic route planning system (Auto-Syn-Route) using quantum computing with tree-tensor network topology with a heuristic scoring function. AutoSynRoute successfully reproduced published synthesis routes for the four case products.

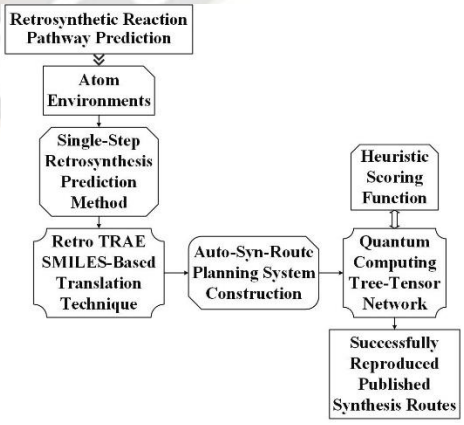

Figure 1. Block Diagram of the Proposed Method

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Figure 1 represents the block diagram of the proposed work. The proposed work consists of a single-step retrosynthetic prediction method. It is predicted using the Retro TRAE SMILES-based translation technique. The SMILES representations of molecular structures are typical inputs for the sequence-to-sequence-based models. However, none of the previously reported models has focused on translation at a sub-structural, fragment, level. Accordingly, a quantum tree-tensor network topology is used for constructing an Auto-Syn-Route planning system.

A. Atom Environments

The concept of circular AEs is employed to represent the molecules in the reaction dataset. Circular environments are defined as topological neighbourhood fragments of varying 'radii' containing all bonds between the included atoms. They are centred on a particular atom, called the central atom. The 'radius' refers to the maximum allowed topological distance between the central atom and all covalently bonded atoms. The topological distance between two atoms was measured as the number of bonds on the shortest path between them. Thus, an AE of radius "r" contains all the atoms in a molecule with a topological distance r or smaller from the central atom and all bonds between them.

1. Single-Step Retrosynthesis

The goal of single-step retrosynthesis is to predict sets of molecules that react to a given product. Since a molecule can be synthesized in various ways, this represents a one-to-many task. Performance in this setting is usually measured by reactant top-k accuracy using a reaction database. This metric measures the fraction of samples for which, given the product of a recorded reaction, the recorded reactants are among the top-k predictions. Given the one-to-many setting, small values of k might not be an optimal choice as there might exist scenarios where a good model receives low scores. Choosing a large k might result in a metric that is overly easy to optimize. Accordingly, the retrosynthesis reaction process is presented in figure 2.

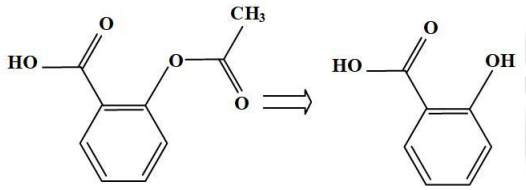

Figure 2. Retrosynthesis Reaction Process

Template-based approaches predict reactant sets via reaction templates. A reaction template encodes atom connectivity changes during a chemical reaction and can be used to transform a product molecule into reactants, $m \xrightarrow{t} r$, where m is a product molecule, r represents a set of reactants and t a reaction template. The product side of a template encodes at which position in a molecule the template can be applied. A necessary condition for this is that the product side of the template is a substructure of the molecule of interest. If this is the case, a template is said to apply to the molecule. The product subgraph is then transformed according to the reactant side of the template and an atom mapping between the two sides. Templates can be either hand-coded or automatically extracted from reaction databases, which yields an ordered set of K unique templates $T = \{t^k\}_{k=1}^{K}$.

Template-relevance prediction aims to predict which templates result in a feasible reaction given a product. If this is the case, a template is relevant to a molecule. While applicability is a necessary condition for relevance, it ignores the context of the whole molecule and thus substructures that might conflict with the encoded reaction. To evaluate template-relevance predictions, use template top-k accuracy, which given the product of a recorded reaction measures the fraction of samples for which the template extracted from the recorded reaction is among the top-k predicted ones.

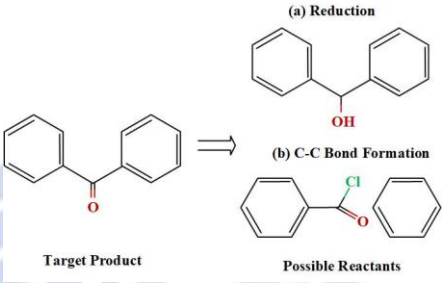

Figure 3. Single-Step Retrosynthesis with Two Possible Reactions

Figure 3 illustrates the single-step retrosynthesis by the Reduction reaction or by the C-C bond formation reaction. Given relevance predictions for a product, reactant sets are obtained by executing top-scoring templates. Do not permit relevance prediction to rely on applicability calculations, because it is relatively slow to compute. Via this constraint, template top-k accuracy also incorporates information about the model's ability to filter out non-applicable templates. This information might be lost in reactant accuracy as template execution relies on a check for applicability. Other differences between the reactant/template accuracy can arise from multiple locations in which the correct template may be applied or incorrect templates leading to the correct reactants. Multistep retrosynthesis can be achieved by applying single-step retrosynthesis recursively. One can decompose the desired molecule into less-complex molecules until only readily available precursor molecules remain.

2. RetroTRAE SMILES-Based Translation Technique

SMILES has been widely used for both forward synthesis prediction and retrosynthesis prediction in the current literature. However, this work, argues that the general-purpose SMILES is deficient for the synthesis prediction problem. Since SMILES is generated by a depth-first traversal of the molecular graph, a molecule can have multiple valid SMILES representations, which leads to the existence of multiple correct output SMILES for a given input SMILES. The one-to-many mapping between input SMILES and output SMILES renders synthesis prediction

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

extremely challenging as the computational model should learn not only the chemical rules for chemical reactions but also the SMILES syntax for SMILES string validity. Several canonicalization methods can be adapted to generate canonical SMILES that ensure a one-to-one mapping between molecules and SMILES. However, these methods are designed for each molecule without considering the relationship between product and reactant molecules, resulting in the large input-output SMILES discrepancy. The large input-output SMILES discrepancy leaves the search space of reactants huge, degrading the performance of synthesis prediction models. Moreover, the canonical SMILES are incompatible with some data augmentation techniques where multiple SMILES are needed for one molecule to bypass the data scarcity issue, as the concept of "canonical SMILES" is violated by multiple SMILES for one molecule.

In this study, a RetroTRAE SMILES-based translation technique is presented. RetroTRAE is starting from a product molecule, it is decomposed into a set of unique integer values. Each AE, a SMART pattern, is associated with a unique integer value. The lists of AEs were provided as input sequences for RetroTRAE. RetroTRAE is trained to predict the proper AE sequences of reactants corresponding to the true reactants. The molecular graph topology is largely unaltered from reactants to products as the molecular changes usually occur locally during chemical reactions. RetroTRAE, using fragment-based tokenization and the Transformer architecture. RetroTRAE mimics chemical reasoning and predicts reactant candidates by learning the changes in atom environments (AEs) associated with the chemical reaction. AEs are the ideal stand-alone chemically meaningful building blocks providing a high-resolution molecular representation. Besides yielding a high level of overall accuracy, the proposed method does not suffer from SMILES-based translation issues such as invalid SMILES. Additionally, the attention matrices of RetroTRAE are shown to capture chemical changes around reaction sites successfully. Through careful inspection of reactant candidates, AEs are promising descriptors for studying reaction route prediction and discovery, which has been underexplored yet.

B. Automatic Data-Driven End-to-End Retrosynthetic Route Planning System

An automatic data-driven end-to-end retrosynthetic route planning system (AutoSynRoute) using quantum computing with tree-tensor network topology with a heuristic scoring function. AutoSynRoute successfully reproduced the published pathways for the four case products, demonstrating its potential for retrosynthetic pathway planning. AutoSynRoute can be applied step-by-step and iteratively with user inputs. To demonstrate this application, predicted the top-10 disconnections for each of the ten reaction classes and reproduced the published retrosynthetic pathways for four examples. To further demonstrate the power of AutoSynRoute, successfully used it to perform automatic retrosynthetic route planning for the above four examples.

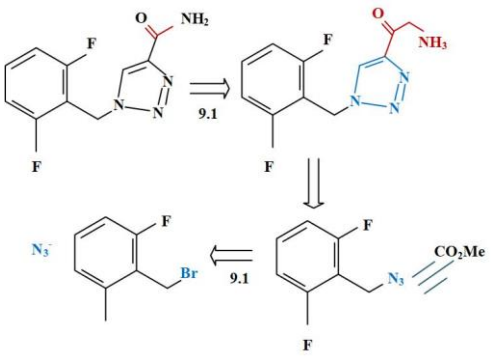

Figure 4. Automatic Retrosynthetic Pathway Planning

Figure 4 represents the automatic retrosynthetic pathway planning [1]. Unlike other template-based methods, which either rely on experts' laborious work or simple, contextless rule-based systems, the approach is fully end-to-end and naturally incorporates the global molecular context of the reaction species. Demonstrated a template-free approach that can be used to perform automatic retrosynthetic route planning and reproduce the published synthesis routes of valuable compounds.

1. Quantum Computing with Tree-Tensor Network Topology

The quantum-computing variants are still a nascent variety, the bulk of the applications to be discussed will involve classical ML algorithms on quantum data even though the focus will certainly be on how the capabilities in each domain can be augmented with the former in the arsenal.

Qubits: The qubit is the basis for all quantum computing, similar to its classical counterpart, the bit. But, there is a significant advantage of the qubit. Unlike the classical bit a qubit stores a mix of two states together, which is called superposition. For a single qubit, the states $|0\rangle = \begin{bmatrix}1\\0\end{bmatrix}$ and $|1\rangle = \begin{bmatrix}0\\1\end{bmatrix}$ are called basis states. It is from these basis states that almost all quantum computation stems.

Quantum Gates: Quantum gates are operations that are performed on qubits, similar to classical gates. These quantum gates are used to change the state of the qubits on which the operation is being performed. They typically are represented in the form of unitary matrices which operate on some initial qubit state. The most common quantum gates are the Hadamard (H), Bit flip (X) and Rotation gate (RX, RY, RZ) which are all single qubit gates. While the Controlled Not (CNOT) is a two-qubit gate. These gates allowed to perform almost all of the basic encodings of data in the quantum state, allowing for meaningful computation of quantum information.

In the gate model of the quantum computing paradigm, transformations between states are achieved using unitary matrices which are represented as 'quantum gates'. Since all quantum gates are unitary, the inverse of such gates necessarily exists and hence transformations using quantum gates alone are always reversible. The way to incorporate irreversibility into the paradigm is through making projective measurements that disturb the state vector irrevocably making it lose its present

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

memory (interactions with the environment induce irreversibility too in the form of qubit decoherence. Then return to this point later). Commonly used quantum gates and their matrix representation in the computational basis. For visualization of the operations of single-qubit gates. However, for $R_n(\theta)$ the axis of rotation n can be either $\{x, y, z\}$ and that decides the accessible state space for a given initial state. For Hadamard transformation, the operation can be viewed as rotation about the axis $(n_x, n_y, n_z)^T = \left(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}}\right)$ through an angle of π and hence creates the state $\frac{|0\rangle+|1\rangle}{2}$ starting from $|0\rangle$. The S-gate $\left(P\left(\frac{\pi}{2}\right)\right)$ and T-gate $\left(P\left(\frac{\pi}{4}\right)\right)$ control the relative phases of $|0\rangle$ and $|1\rangle$. These operations are commonly used to entangle two or more qubits in a quantum circuit.

Tree Tensor Networks (TTNs)

Tree tensor networks provide another approach to model quantum states by arranging the local tensors in a tree-like pattern. A TTN can be formed from an n-qubit quantum state using the tree-Tucker decomposition. Like other tensor networks, TTNs are used as an ansatz to simulate the ground state of the local Hamiltonian. Tensors in TTNs form the nodes of the tree which are connected through bond indices. The physical indices appear on the leaf nodes. On contracting the bond indices, the TTN has n free indices which represent the physical degree of freedom of the state. TTNs are a generalization of MPS and can in principle be non-binary as well. An MPS can be thought of as a flattened TTN such that each parent node has one successor (bond indices of MPS) and another leaf node (physical indices of MPS).

The structure of TTN is inspired by the spatial renormalization group. At every layer of TTN, coarse-graining is carried out between neighbouring sub-trees. Unlike MPS, the local tensors with access to physical indices in TTN are not connected directly to each other, the correlation between qubits is represented through the layers. The local correlation information is stored in the lower layers while the upper layers store long-range correlation information.

Each node in a TTN is a three-dimensional tensor (except the root/uppermost node) with at most one upper index α and two lower indices β_1 and β_2. The tensors can be written as $w^{\alpha}_{\beta_1\beta_2}$. The space required to store a TTN grows as $O(ND^3)$, where N is the number of physical indices and D is the bond dimension of the local tensors. Each tensor in TTN is an isometry satisfying the following condition:

$$\sum_{\beta_1,\beta_2}(w)^{\alpha}_{\beta_1,\beta_2}(w^*)^{\beta_1,\beta_2}_{\alpha} \quad (1)$$

Choosing an isometric tensor as in equation (1) is advantageous in numerous ways. It simplifies the optimization of TTN and calculation of the expectation values of local observables and it is also known to provide numerical stability to TTN algorithms. TTN can very well be generalized to higher dimensions by appropriately placing isometries across local physical indices and hierarchically merging sub-trees through more isometries.

Tree tensor networks form the basis of the multi-layer multi-configuration time-dependent Hartree methods which are used to perform quantum molecular dynamics simulations.

Evaluation Procedure

To evaluate the performance of the translation model, a suitable metric was required to measure the similarity between predictions and the true reactants. The Tanimoto (T_c) and the Sørrensen-Dice coefficient (S) as two of the special cases of the Tversky index were the similarity metrics used in this study. The exact form of the Tversky index is as follows:

$$S(X,Y) = \frac{|X \cap Y|}{|X \cap Y| + \alpha|X - Y| + \beta|Y - X|} \quad (2)$$

Here, α, β ≥ 0 are the parameters of the Tversky index. Setting α = β = 1 leads to the Tanimoto coefficient; setting α = β = 0.5 leads to the Sørrensen-Dice coefficient. The Tanimoto and Dice coefficients measured between two molecules range between 0 and 1. The value of zero represents the total dissimilarity, whereas a value of 1 represents the exact match.

Unlike SMILES-based methods, small prediction errors of the AE representation do not yield invalid predictions. Thus, multiple degrees of accuracy can be calculated due to the native design of this model. The results were computed with four different cutoffs, which can be categorized as (a) hard thresholds, and (b) soft thresholds. Here, define hard thresholds as the discrepancies of one or two fragments. Call arbitrary thresholds based on the Tanimoto coefficient soft thresholds such as $T_c \geq 0.85$. These measures are conventionally used to screen similar molecules. For example, molecules having $T_c \geq 0.85$ tend to exhibit similar biological activities.

Hard thresholds offer the following advantages over soft thresholds. First, hard thresholds do not depend on sequence length. Second, contrary to soft thresholds, easily find the type and number of fragments that deviated from the ground truth. Finally, by using hard thresholds, any risk of losing high-quality reactant candidates can be avoided which could be excluded with soft thresholds. This suggests that high-quality predictions with low and medium complexity, relatively smaller molecules, have a higher chance of being excluded by soft thresholds. For example, a high-quality double mutated prediction with medium complexity represented with 13 AEs could be overlooked by a bio-actively similar threshold $(T_c \geq 0.85)$.

In this study, a top-1 prediction is used as the best recommendation to report the performance of the model, as well as for molecular search and retrieval. Since there are many ways to decompose a molecule, retrosynthetic prediction tools can procure many different possible synthetic routes. However, the analyses showed that only 6% of the USPTO dataset has at least two sets of reactants. Thus, using top-1 accuracy is a legitimate measure to assess a single-step retrosynthesis predictor trained on the USPTO dataset. Top-N accuracy for evaluating retrosynthesis prediction has recently been disputed because, with each prediction, a model tends to find the next frequently observed answer among reactions in a dataset rather than making a chemically more meaningful prediction. A few alternative metrics were newly suggested, such as Round-trip, and MaxFrag.

V. Experimentation And Result Discussion

The transformer connects the encoder and decoder units to translate between sequences by effectively employing a multi-head attention mechanism on each unit. Input and output sequences for the proposed model are the lists of AEs. Accordingly, tested several different schemes to convert

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

molecules into a list of fragments, such as MACCS keys, and AEs. AEs are fragments consisting of a central atom and its covalently bonded neighbours with a predefined radius. They can be considered the basis of constructing molecules, similar to the pieces of a jigsaw puzzle.

TABLE I. TABLE OF SYSTEM CONFIGURATION FOR SIMULATION

Simulation System Configuration	
Python Jupiter	Version 3.8.0
Operation System	Ubuntu
Memory Capacity	4GB DDR3
Processor	Intel Core i5 @ 3.5GHz
Simulation Time	50 seconds

The proposed methods are analysed using Python software with version 3.8.0. The operating system of the proposed work is Ubuntu and its memory capacity is 4GB DDR3. The processor of the implemented work is an Intel core i5 @ 3.5 GHz, and their simulation time is 50 seconds, these details are presented in table 1.

Performance of RetroTRAE

Prediction performance, as a function of different similarity thresholds for RetroTRAE. RetroTRAE has reached top-1 exact match accuracies of 67% trained with 10 times augmented uni- and bimolecular datasets. Augmentation slightly improved the results and stabilized the model's learning since more data and randomness were added to the network. Although the AE representation is permutation invariant, the models with positional encoding perform better than those trained without using positional information.

One of the advantages of using AEs over SMILES is that a few errors do not lead to invalid predictions. Thus investigated how much the success rate can be improved by easing the threshold without losing the functionality of the retrosynthetic framework. When single mutations (SM) were allowed, the success rates of uni-molecular and bi-molecular reactions increased to 58.1 and 60.9%, respectively. The corresponding numbers for double mutations (DM) were 60.5 and 62.7%. To quantify how low the probability of finding such extremely close neighbours of molecules is in a large database, performed extensive analysis by using AEs. Considering the cumulative distribution function of AEs obtained with 1.3 million molecules in the USPTO database, only 13 pairs were found to have a T_c value of 0.76 or higher. With a threshold of 0.9 or higher, most molecules in a typical database would be singletons with no near neighbours.

The mean T_c of all predictions of the uni-molecular test set was found to be 0.88, which is highly statistically significant with a p -value $< 10^{-5}$. This indicates that even non-exact predictions made by RetroTRAE are still highly similar to the ground truth.

Model Interpretability

It is often difficult to attribute meaning to the outcomes of deep learning methodologies. Accordingly, identified that the proposed model successfully learned the changes in chemical environments around reaction centres. In contrast to this work, in SMILES-to-SMILES translations chemical changes mostly occur via rearrangements of SMILES tokens rather than actual transformations of chemically meaningful tokens, which hampers chemical interpretability and explainability.

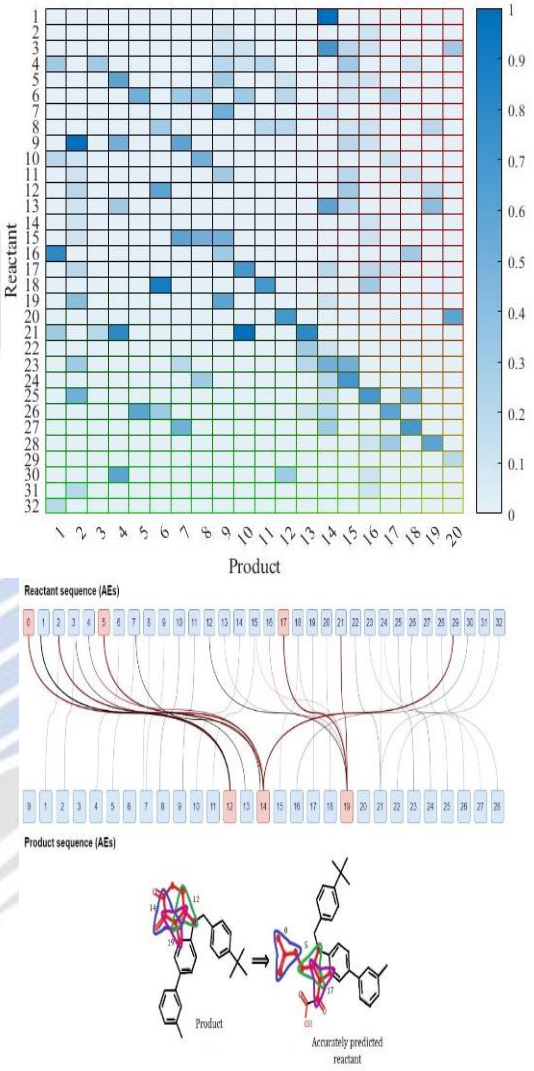

Figure 5. Interpretability of RetroTRAE for Uni-Molecular Ring-Opening Reaction

The attention weight matrices and the fragments with the highest attention values of uni-molecular ring-opening reaction are visualized in figure 5. The AE that changes the reaction has the highest attention value with its changed counterpart. Likewise, the AEs that remain intact tend to have the highest attention. The column-wise summations of attention weights indicate the mostly attended AEs of a product by RetroTRAE. To show this, the AEs in products that changed during the reactions and their attention on the reactant side is highlighted. Indeed, the model pays more attention to altered AEs near the reaction centres as exemplified by ring opening and dissociation reactions. These examples clearly show that AE tokens are chemically meaningful and fully interpretable by themselves as opposed to SMILES tokens.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

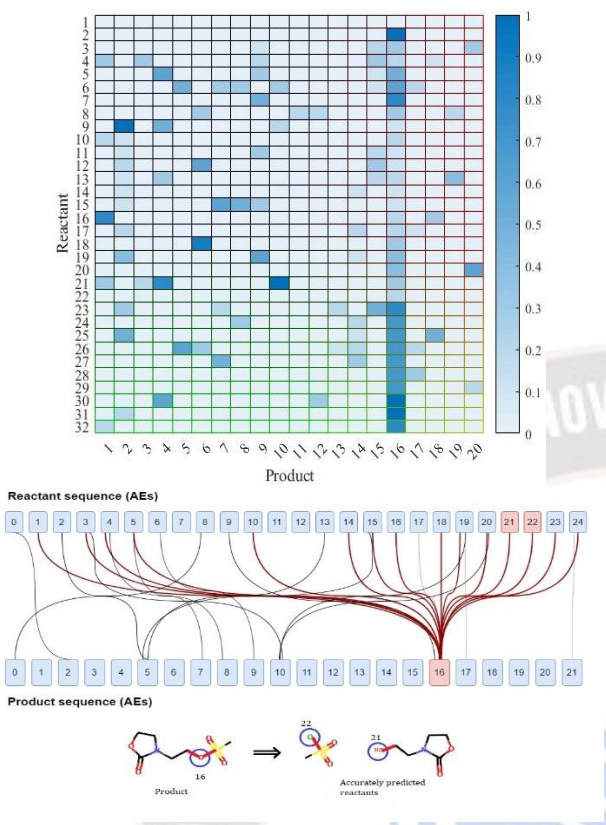

Figure 6. Interpretability of RetroTRAE for Bi-Molecular Dissociation Reaction

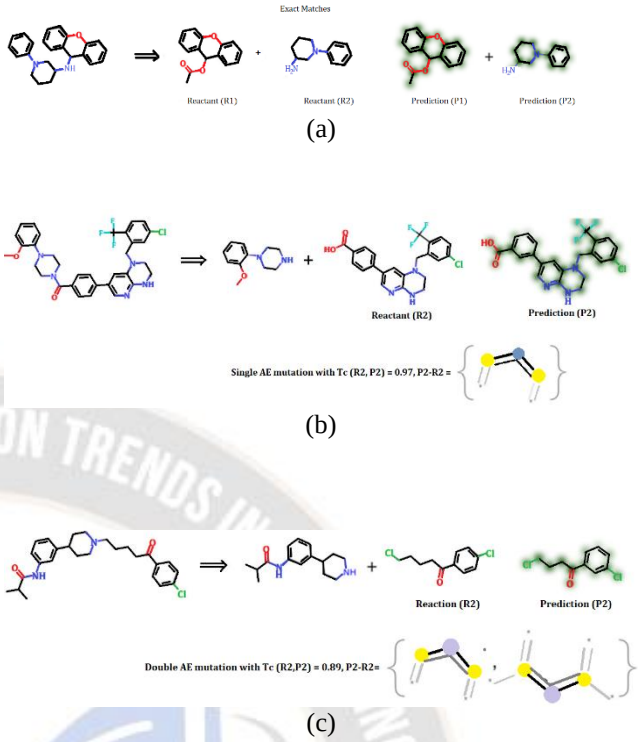

Figure 7. Exact RetroTRAE Predictions

Figure 6 portrays the attention weight matrices and the fragments with the highest attention values of Bi-molecular dissociation reaction. RetroTRAE operates at the level of AEs predicting transformations from products to reactants in a single step. The main reason for focusing on single-step reactions is that the mechanistic descriptions of reactions are not provided in the USPTO database. However, there is no intrinsic limitation for the model to predict multi-step synthetic routes. The model would be able to predict multi-step synthetic routes when it is combined with a proper algorithm. In its current form, RetroTRAE can be used in any single step of a multi-step retrosynthesis.

Retrosynthesis Predictions

The candidate reactants with $T_c > 0.85$ are similar enough to their true counterparts. To validate this assumption, assessed the quality of candidate reactants by comparing them with true reactants. The accuracy of side-substituents is regarded as less significant for matching the reactants' functionality, especially when they are simple alkyls. In addition to exact predictions, investigated how much singly and doubly mutated predictions are similar to the ground truth.

The exact prediction based on RetroTRAE is presented in figure 7 (a). RetroTRAE predicted 58.1% of the reactions in the test set accurately. The single and double fragment mutations together account for 3.3% of the total predictions. In single mutation cases, atom and connectivity types must be preserved, therefore only two types of structural changes are possible. First, a new environment may appear (or an existing environment may disappear) due to a misplaced single environment (e.g., at the ortho/para/meta position). With this change, all connected atom types must be preserved in figure 7 (b) which is a hard threshold. Second, a single existing AE can be added or subtracted at terminal sites. Double mutations are characterized by a misplaced branching AE or a single atom substitution shown in figure 7 (c). If a mutation happens in the middle of a molecule, the AE centred at the mutated site and its direct neighbours are highly likely to be changed, leading to at least three AE mutations.

As indicated in the similarity maps of hard thresholds, none of the atoms of the reactant candidates negatively contributed (red) to the similarity value. With the AE representation, the length of simple aliphatic chains might be incorrectly predicted, because the length of an aliphatic chain cannot be accurately described using a set of unique fragments. Based on this observation, SM and DM predictions are much more similar to the ground truth than conventional structural analogues implying differences in certain substructures, functional groups, or several atom types. We believe that these small discrepancies are easily amendable through a visual comparison with a product. When soft thresholds are used, several AEs can be

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

altered, making the generalization of errors highly difficult. After inspecting the bio-actively similar predictions, the most significant aspects of the retrosynthetic analysis concluded, such as bond disconnections, reactive functional groups, and core structures, were correctly predicted. Nevertheless, unable to generalize the characteristics of the predictions beyond DMs, albeit within the bounds of bioactive similarity space.

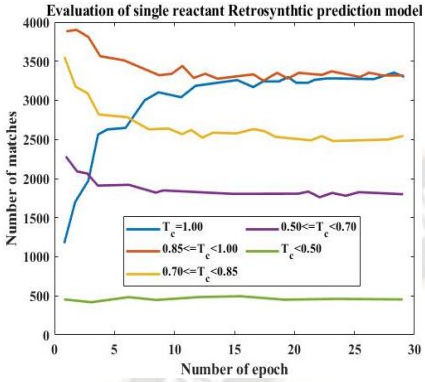

Figure 8. Number of Matches at Different Numbers of Epochs

The evolution of prediction accuracy concerning threshold values along the training epoch for the single reactant validation set is illustrated in figure 8. In particular, it is demonstrated that the network successfully learned reaction rules by capturing the alterations of molecules at a sub-structural level. The number of exact matches ($T_c = 1.0$) increased rapidly during the first 10 epochs. After 20 epochs, the value became almost tripled. The likelihood of making a better prediction for each fragment becomes higher during training. This is a clear indication of successful training. The improvement in exact matches appears to be a result of the respective declines in non-exact matches except extremely bad predictions ($T_c < 0.5$). The quality of bad predictions did not improve probably due to the insufficient information, complexity, and noise contained in the data. This observation was similarly repeated for all the other datasets.

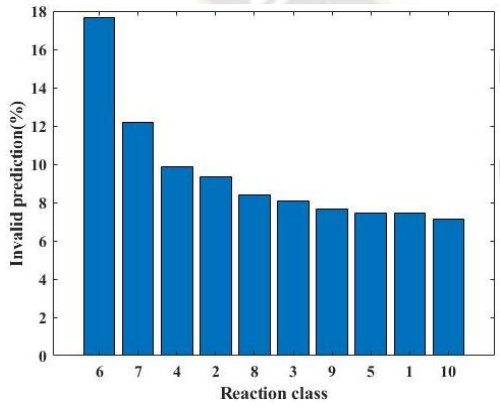

Figure 9. Invalid Predictions for Top 10 Predictions with Known Reaction Classes

The fraction of invalid predictions across the various known reaction types for top-10 analysis is presented in figure 9. To further understand the performance of the proposed model across reaction classes, the granularity of the analysis and compute the five metrics– accuracy, fractional accuracy, MaxFrag accuracy, similarity score, and syntactic validity across the 10 reaction classes are increased. This behaviour is not trivial since the corresponding top-10 prediction accuracy does not follow the same trend across reaction classes. Moreover, the percentage of invalid predictions shows only minor variations across the two scenarios with known and unknown reaction classes. This observation again highlights the ability of this proposed Retro TRAE SMILES-based translation technique generates correct predictions, irrespective of the other factors.

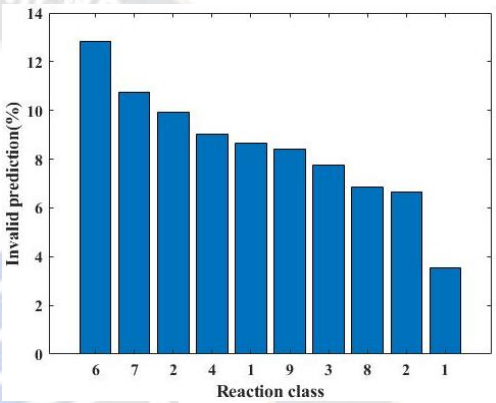

Figure 10. Invalid Predictions for Top-10 Predictions without Reaction Class

Figure 10 reveals the invalid percentages for Top-10 predictions without reaction class. The above trend indicates that except for reaction class 6 (deprotections) and the surprisingly accurate predictions on reaction class 10 (functional group addition) when the reaction class is unknown, all the reaction types result in nearly the same percentage of invalid predictions. The high percentage error in deprotection reactions could be attributed to several factors that could be specific to the reaction class and could be analysed through chemistry-driven heuristics that envision as a hybrid explanation-generation system, respectively.

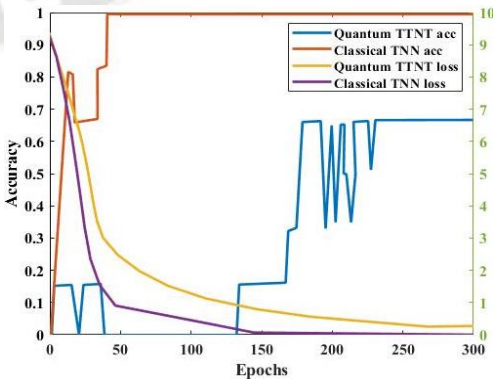

Figure 11. Results of Chemical Retrosynthesis

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

The chemical retrosynthesis results based on training the small amount of chemical sample limited data set of a single reaction type (reaction type 1) is presented in figure 11, where quantum can reach reasonable accuracy. Here, train the quantum and tree-tensor network topology. The promising results in this figure show that the quantum approach, while unable to match the results of the classical approach, can converge to an accuracy of 68% and a loss of 0.1, respectively.

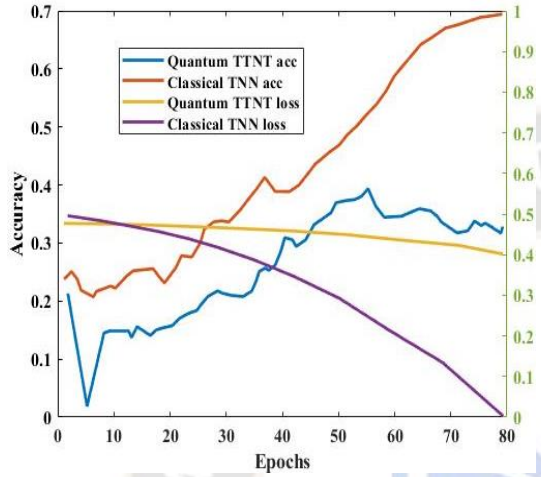

Figure 12. Chemical Retrosynthesis Results for Training Huge Number of Chemical Samples

The results in figure 12 show that the classical loss never reaches a point of convergence, whereas the quantum loss also doesn't reach convergence nor does it reach the same level as the classical. These results hold for accuracy, where the classical domain reaches 65% and the quantum domain reaches 55%. While there is a small performance gap, the task of identifying a common substring within the predicted reactants, and quantum can nearly match classical performance during training. Validation is run once every 5 epochs during training and here, there is a flip of performance.

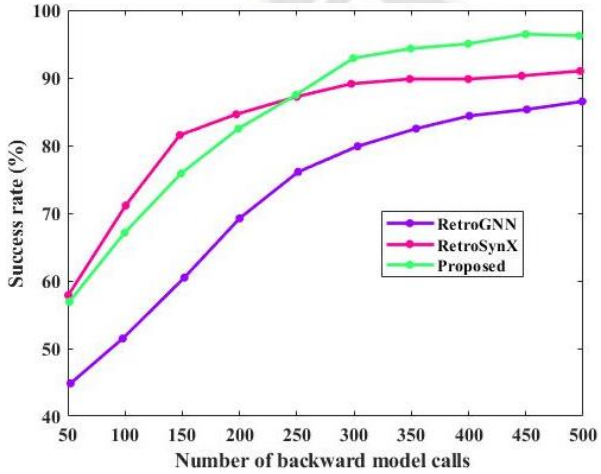

Figure 13. Success Rate under Reaction Model

The results for evaluating the performance of the proposed framework and baselines are in figure 13. Compare the framework with existing RetroGNN [26] and RetroSynx [19]. While compare to these existing methods, the proposed work significantly outperforms the RetroGNN and RetroSynx. However, the proposed Retro-TRAE SMILES achieves a success rate of 97% with a computation limit of $N = 500$. While the other two methods like RetroSynx and RetroGNN achieve 90.65% and 86%. In terms of other evaluation metrics, e.g., the length and the cost of discovered reaction pathways. Therefore, such a result demonstrates the effectiveness of the proposed framework for retrosynthetic planning.

On the other hand, the backward reaction model does not suffer from a drop in TOP-k accuracy. Instead, they even outperform the backward reaction model trained on the reaction dataset $D_{reaction}$ in terms of TOP-10 accuracy. An improvement in the TOP-10 accuracy comes from "diverse" solution candidates generated by the proposed model, which is encouraged by being trained on a large variety of samples, e.g., augmented reactions. The performance of the proposed method may be saturated when the backward reaction model appropriately adapts to the search algorithm. This highlights the importance of training an appropriate backward reaction model for retrosynthetic planning.

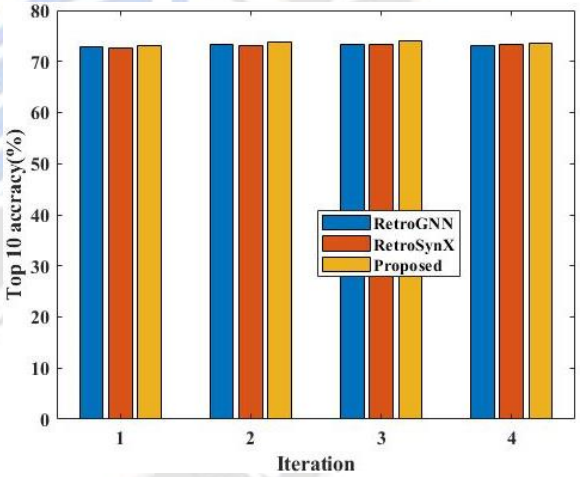

Figure 14. Top-10 Accuracy under Multiple Iteration

Figure 14 reveals the top 10 accuracies for numerous iterations. The performance of the proposed model for finding the reaction pathways over multiple iterations. The result demonstrates that iterating the framework improves the success rate of finding a reaction pathway while maintaining the accuracy of the backward reaction model. This indeed validates that the proposed framework is effective without compromising the ability to model realistic reactions. The proposed method is compared with the existing RetroGNN [26] and RetroSynx [19], when compared to these methods, the proposed work produces higher accuracy.

VI. Research Conclusion

Retrosynthesis analysis is a challenging problem since it involves predicting the precursors with limited information,

searching a combinatorially large number of possible synthesis pathways, and approximating an often complex multi-step analysis as a single-step prediction problem. Naturally, incorporating additional information about the reaction or the molecules involved would be of considerable use given the complexity of the task and the limited information often present for making the predictions. In this study, RetroTRAE SMILES-based translation technique is proposed for a single-step retrosynthetic prediction by associating the AEs of the reactants with the products. Throughout the study, AEs are regarded as the basis of molecules and employed in the prediction workflow. The proposed design enables to capture of chemical changes by focusing on fragments related to the reaction centres. Subsequently, the research presents a quantum tree-tensor network topology with a heuristic scoring function for constructing an Auto-Syn-Route planning system. The AutoSynRoute successfully reproduced published synthesis routes for the four case products. The system is analyzed on a Python software platform.

- ❖ To accurately generate the reactant candidates for a target molecule, the proposed model is used. Therefore, the proposed work showed that the proposed model achieves a top-1 exact matching accuracy of 68%.
- ❖ The overall accuracy increased to 73% by adding extremely similar predictions. These results are better than those of the existing methods, without suffering from problems associated with the SMILES representation.
- ❖ RetroTRAE showed comparable or improved performance compared to other state-of-the-art models like RetroGNN and RetroSynx.
- ❖ The retrieval process converts a set of fragments into a molecule concerning coverage, degeneracy, and resolution. RetroTRAE predicted reactant candidates with an exact match accuracy of 73%.

In addition to the exact match accuracy, highly similar reactant candidates with single and double mutations were exceptionally similar to ground truth. The overall accuracy with singly and doubly mutated predictions was 73%, therefore, the proposed research work outperforms the state-of-the-art methods. The proposed findings will open new possibilities for the development of different models for chemistry using sequential data, not only for retrosynthetic prediction but also for reaction and property predictions, respectively.

REFERENCES

[1] K. Lin, Y. Xu, J. Pei, and L. Lai, "Automatic retrosynthetic route planning using template-free models," Chemical science, vol. 11, no. 12, pp.3355-3364, 2020.
[2] P. Schwaller, R. Petraglia, V. Zullo, V.H. Nair, R.A. Haeuselmann, R. Pisoni, C. Bekas, A. Iuliano, and T. Laino, "Predicting retrosynthetic pathways using transformer-based models and a hyper-graph exploration strategy," Chemical science, vol. 11, no. 12, pp.3316-3325, 2020.
[3] S. Chen, and Y. Jung, "Deep retrosynthetic reaction prediction using local reactivity and global attention," JACS Au, vol. 1, no. 10, pp.1612-1620, 2021.
[4] S.A. Vajiha Begum, M. Pushpa Rani, "Recognition of neurodegenerative diseases with gait patterns using double feature extraction methods," IEEE - 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 332-338, 2020, 10.1109/ICICCS48265.2020.9120920
[5] Nihar Ranjan Kar, "Osmotically Controlled Drug Delivery Systems: A Review," Indian Journal of Natural Sciences, vol. 13, no. 72, pp. 43372-43394, 2022.
[6] K. Lin, Y. Xu, J. Pei, and L. Lai, "Automatic retrosynthetic pathway planning using template-free models," arXiv preprint arXiv:1906.02308, 2019.
[7] Y. Mo, Y. Guan, P. Verma, J. Guo, M.E. Fortunato, Z. Lu, C.W. Coley, and K.F. Jensen, "Evaluating and clustering retrosynthesis pathways with learned strategy," Chemical science, vol. 12, no. 4, pp.1469-1478, 2021.
[8] Y. Wang, C. Pang, Y. Wang, Y. Jiang, J. Jin, S. Liang, Q. Zou, and L. Wei, "MechRetro is a chemical-mechanism-driven graph learning framework for interpretable retrosynthesis prediction and pathway planning," arXiv preprint arXiv:2210.02630, 2022.
[9] Athi sakthi, Pushpa Rani, "Detection of Movement Disorders Using Multi SVM," Global Journal of Computer Science and Technology, vol. 13, no.1, pp. 23-25, 2013.
[10] Nihar Ranjan Kar, "Nanotechnology-Based Targeted Drug Delivery Systems for Brain Tumors," Indian Journal of Natural Sciences, vol. 12, no. 68, pp. 35081-35094, 2021.
[11] A. Thakkar, V. Chadimová, E.J. Bjerrum, O. Engkvist, and J.L. Reymond, "Retrosynthetic accessibility score (RAscore)–rapid machine learned synthesizability classification from AI driven retrosynthetic planning," Chemical science, vol. 12, no. 9, pp.3339-3349, 2021.
[12] S. Genheden, and E. Bjerrum, "PaRoutes: a framework for benchmarking retrosynthesis route predictions," 2022.
[13] J.L. Baylon, N.A. Cilfone, J.R. Gulcher, and T.W. Chittenden, "Enhancing retrosynthetic reaction prediction with deep learning using multiscale reaction classification," Journal of chemical information and modeling, vol. 59, no. 2, pp.673-688, 2019.
[14] A. Komathi, M. Pushpa Rani, "Trust performance of AODV, DSR and DSDV in wireless sensor networks," IEEE - Second International Conference on Current Trends In Engineering and Technology," pp. 423-425, 2014.
[15] Nihar Ranjan Kar, "Liposomal Drug Delivery System: An Overview," Shodhasamhita: Journal of Fundamental & Comparative Research, vol. VIII, no. 5, pp. 123-134, 2022.
[16] S. Genheden, O. Engkvist, and E. Bjerrum, "Fast prediction of distances between synthetic routes with deep learning," Machine Learning: Science and Technology, vol. 3, no. 1, pp.015018, 2022.
[17] B. Chen, C. Li, H. Dai, and L. Song, "Retro*: learning retrosynthetic planning with neural guided A* search," In International Conference on Machine Learning, pp. 1608-1616, 2020, November. PMLR.
[18] M.E. Fortunato, C.W. Coley, B.C. Barnes, and K.F. Jensen, "Data augmentation and pretraining for template-based retrosynthetic prediction in computer-aided synthesis planning," Journal of chemical information and modeling, vol. 60, no. 7, pp.3398-3407, 2020.
[19] A. Komathi, M. Pushpa Rani, "Shift reduce parser based malicious sensor detection for predicting forest fire in WSNs," Wireless Personal Communications, vol. 103, pp. 2843-2861, 2018.
[20] K. Lin, Y. Xu, J. Pei, and L. Lai, "Automatic retrosynthetic route planning using template-free models," Chemical science, vol. 11, no. 12, pp.3355-3364, 2020.
[21] Nihar Ranjan Kar, "A Review on Targeting of Drugs to Brain," Indian Journal of Natural Sciences, vol. 11, no. 64, 2021.
[22] M. Zhao, L. Fang, L. Tan, J.G. Lou, and Y. Lepage, "Leveraging Reaction-aware Substructures for Retrosynthesis and Reaction Prediction," arXiv preprint arXiv:2204.05919, 2022.
[23] N. Lee, J. Jeong, and D. Shin, "AI System for Supporting Generation of Optimal Synthetic Pathways Based on Chemical Reaction Big Data," In Computer Aided Chemical Engineering, vol. 50, pp. 1053-1058, 2021. Elsevier.
[24] N. Mary Peter, M, Pushpa Rani, "V2V communication and authentication: the internet of things vehicles (IoTV),"

Wireless Personal Communications, vol.120, no.1, pp. 231-247, 2021.
[25] U.V. Ucak, I. Ashyrmamatov, J. Ko, and J. Lee, "Retrosynthetic reaction pathway prediction through neural machine translation of atomic environments," Nature communications, vol. 13, no. 1, pp.1-10, 2022.
[26] J. Jeong, N. Lee, Y. Shin, and D. Shin, "Intelligent generation of optimal synthetic pathways based on knowledge graph inference and retrosynthetic predictions using reaction big data," Journal of the Taiwan Institute of Chemical Engineers, vol. 130, pp.103982, 2022.
[27] S. Zheng, J. Rao, Z. Zhang, J. Xu, & Y. Yang, "Predicting retrosynthetic reactions using self-corrected transformer neural networks," Journal of chemical information and modeling, vol. 60, no. 1, pp. 47-55, 2019.
[28] P. Brinda, M. Pushpa Rani, "Analysis of Early Leave Pest Detection," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 4, no.5, pp. 2321-8169, 2016.
[29] W. Wang, Q. Liu, L. Zhang, Y. Dong, and J. Du, "RetroSynX: A retrosynthetic analysis framework using hybrid reaction templates and group contribution-based thermodynamic models," Chemical Engineering Science, vol. 248, pp.117208, 2022.
[30] X. Wang, Y. Li, J. Qiu, G. Chen, H. Liu, B. Liao… & X. Yao, "Retroprime: A diverse, plausible and transformer-based method for single-step retrosynthesis predictions," Chemical Engineering Journal, vol. 420, pp. 129845, 2021.
[31] W. Wang, Q. Liu, L. Zhang, Y. Dong, and J. Du, "Retrosynthesis Pathway Design Using Hybrid Reaction Templates and Group Contribution-Based Thermodynamic Models," In Computer Aided Chemical Engineering, vol. 49, pp. 85- 90, 2022. Elsevier.
[32] S. Zheng, T. Zeng, C. Li, B. Chen, C.W. Coley, Y. Yang, and R. Wu, "Deep learning driven biosynthetic pathways navigation for natural products with BioNavi-NP," Nature Communications, vol. 13, no. 1, pp.1-9, 2022.
[33] H. He, J. Wang, Y. Liu, and F. Wu, "Modeling Diverse Chemical Reactions for Single-step Retrosynthesis via Discrete Latent Variables," arXiv preprint arXiv:2208.05482, 2022.
[34] S. Ishida, T K. erayama, R. Kojima, K. Takasu, and Y. Okuno, "AI-Driven Synthetic Route Design Incorporated with Retrosynthesis Knowledge," Journal of chemical information and modeling, vol. 62, no. 6, pp.1357-1367, 2022.
[35] Y. Jiang, Y. Yu, M. Kong, Y. Mei, L. Yuan, Z. Huang, K. Kuang, Z. Wang, H. Yao, J. Zou, and C.W. Coley, "Artificial Intelligence for Retrosynthesis Prediction," Engineering, 2022.
[36] C.H. Liu, M. Korablyov, S. Jastrzębski, P. Włodarczyk-Pruszyński, Y. Bengio, and M. Segler, "RetroGNN: Fast Estimation of Synthesizability for Virtual Screening and De Novo Design by Learning from Slow Retrosynthesis Software," Journal of Chemical Information and Modeling, vol. 62, no. 10, pp.2293-2300, 2022.
[37] A.P. Senthil Kumar, S. Yuvaraj, S. Janaki, "Experimental investigations of Co3O4, SiO2, cotton seed oil additive blends in the diesel engine and optimization by ANN-SVM process," Journal of Ceramic Processing Research, vol. 21, no. 2, pp. 217-225, 2020.
[38] S. Yuvaraj, A.P. Senthil Kumar, M. Muthukumar, K. Sadesh, S. Janaki, "Certain studies on influence of nano catalysts Co3O4, SiO2 blended with CME-diesel in combustion," Materials Today: Proceedings, vol. 51, pp. 1612-1618, 2022.
[39] Tarun Kumar Kotteda, Manoj Kumar, Pramod Kumar, Rama Bhadri Raju Chekuri, "Metal matrix nanocomposites: future scope in the fabrication and machining techniques," The International Journal of Advanced Manufacturing Technology, 2022. https://doi.org/10.1007/s00170-022-09847-0
[40] Tarun Kumar Kotteda, Rama Bhadri Raju Chekuri, Naga Raju, Prasada Raju Kantheti, S. Balakumar, "Analysis on Emissions and Performance of Ceramic Coated Diesel Engine Fueled with Novel Blends Using Artificial Intelligence," Advances in Materials Science and Engineering, 2021, https://doi.org/10.1155/2021/7954488
[41] R.K. Pattanaik, S. Mishra, M. Siddique, T. Gopikrishna, S. Satapathy, "Breast Cancer Classification from Mammogram Images Using Extreme Learning Machine-Based DenseNet121 Model," Genetics Research, pp. 2731364, 2022.
[42] S.K. Mohapatra, S. Prasad, G.M. Habtemariam, M. Siddique, "Identify determinants of infant and child mortality based using machine learning: Case study on Ethiopia," Big Data Analytics and Machine Intelligence in Biomedical and Health Informatics: Concepts, Methodologies, Tools and Applications, pp. 21–45, 2022.
[43] M. Siddique, D. Panda, "Prediction of stock index of tata steel using hybrid machine learning based optimization techniques," International Journal of Recent Technology and Engineering, vol. 8, no. 2, pp. 3186–3193, 2019.
[44] M. Siddique, D. Panda, "A hybrid forecasting model for prediction of stock index of tata motors using principal component analysis, support vector regression and particle swarm optimization," International Journal of Engineering and Advanced Technology, vol. 9, no. 1, pp. 3032–3037, 2019.
[45] S. S. Reddy, N. Sethi, R. Rajender, & V.S.R. Vetukuri, "Non-invasive diagnosis of diabetes using chaotic features and genetic learning," In Third International Conference on Image Processing and Capsule Networks, pp. 161–170, 2022. Cham: Springer International Publishing.
[46] S. S. Reddy, L. Alluri, M. Gadiraju, & R. Devareddi, "Forecasting Diabetic Foot Ulcers Using Deep Learning Models," Proceedings of Third International Conference on Sustainable Expert Systems, pp 211–227, 2023.
[47] R. S. Shankar, D. R. Babu, K.V.S.S. Murthy, & V. Gupta, 'An approach for essay evaluation using system tools," 2017 International Conference on Innovative Research In Electrical Sciences (IICIRES). IEEE. 2017.
[48] R. Shiva Shankar, & D. Ravibabu, "Digital report grading using NLP feature selection," In Soft Computing in Data Analytics,615–623, 2019. Proceedings of International Conference on SCDA 2018.
[49] Reddy, Shiva Shankar, M. Gadiraju, & V.V.R. Maheswara Rao, "Analyzing student reviews on teacher performance using long short-term memory," In Innovative Data Communication Technologies and Application, pp. 539–553, 2022. Singapore: Springer Nature Singapore.