**A HYBRID QUANTUM-CLASSICAL FUSION NEURAL NETWORK TO IMPROVE**

**PROTEIN-LIGAND BINDING AFFINITY PREDICTIONS FOR DRUG DISCOVERY**

L. Domingo1⋆ , M. Chehimi1 , S. Banerjee2 , S. He Yuxun2 , S. Konakanchi2 ,

L. Ogunfowora2 , S. Roy2 , S. Selvarajan2 , M. Djukic1 and C. Johnson1⋆

2 Ingenii Inc., New York, USA,

1 Purdue University, The Data Mine, USA ,

*⋆*Corresponding Author’ Emails: [laia@ingenii.dev](mailto:laia@ingenii.dev)

|  |  |
| --- | --- |
| The field of drug discovery hinges on the accurate predic  tion of binding affinity between prospective drug molecules  and target proteins, especially when such proteins directly  influence disease progression. However, estimating binding  affinity demands significant financial and computational re  sources. While state-of-the-art methodologies employ classi  cal machine learning techniques, emerging hybrid quantum  machine learning models have shown promise for enhanced  performance, owing to their inherent parallelism and capac  ity to manage exponential increases in data dimensionality.  Despite these advances, existing models encounter issues  related to convergence stability and prediction accuracy.  This paper introduces a novel hybrid quantum-classical deep  learning model tailored for binding affinity prediction in  drug discovery. Specifically, the proposed model synergis  tically integrates 3D and spatial graph convolutional neural  networks within an optimized quantum architecture. Simu  lation results demonstrate a 6% improvement in prediction  accuracy relative to existing classical models, as well as a  significantly more stable convergence performance compared  to previous classical approaches. Moreover, to deploy the  proposed framework over today’s *noisy intermediate-scale*  *quantum (NISQ)* devices, a novel quantum error mitigation  algorithm is proposed. This algorithm outperforms existing  techniques and provides optimal error mitigation for error  probabilities *p ≤* 0*.*05, while resulting in no additional  overhead during the training and testing phases.  ***Index Terms*—Quantum machine learning (QML), drug**  **discovery, binding affinity, quantum fusion model.**  **I. INTRODUCTION**  The healthcare landscape has undergone a transformative  shift, notably marked by advancements in drug discov  ery through the integration of emerging technologies with  conventional techniques. Through complex molecular inter  actions and precise computational modeling of compound  interactions, novel drugs are rigorously designed and identi  fied. Central to this drugs design process is the understanding  of proteins and their role in disease mechanisms [1].  All Purdue authors contributed equally.  the work in [9] leveraged 3D-convolutional neural networks  (3D-CNNs) to perform protein-ligand binding affinity pre  dictions in a faster and more efficient manner relative to other  ML models. Moreover, the work in [10] enhanced the model  proposed in [9] by predicting binding affinities using an  ensemble of several independently-trained 3D-CNN network  layers. Furthermore, the work in [11] introduced a classical  fusion model combining a 3D-CNN and a spatial graph CNN  (SG-CNN). The model in [11] enhances the binding affinity  prediction accuracy by concurrently processing grid-based,  context-based, and graph-based protein features. However,  the models proposed in [9]–[11] do not have a stable  convergence, and their predictive accuracy is not optimized.  Additionally, the work in [12] uses quantum support vec  tor machines for virtual drug screening. Although the hybrid  QML model in [12] outperforms classical counterparts, its  focus on a limited subset of protein-ligand features restricts  the model accuracy and scalability to larger datasets. Finally,  the work in [13] proposes a hybrid QML architecture, modi  fying a classical CNN by replacing a classical convolutional  layer with an optimized quantum circuit. While the work  in [13] successfully tackles the computational complexity  inherent in classical neural networks, it does not yield an  improvement in binding affinity predictive accuracy, thus  achieving a comparable performance to classical models. On  top of that, prior works [12] and [13] did not analyze how  the presence of noise in practical NISQ devices impacts  the performance of the quantum solution. To the best of  our knowledge, no existing research has effectively tackled  the issue of binding affinity prediction in drug discovery  by capitalizing on the benefits of QML while simultane  ously achieving high accuracy, ensuring smooth and stable  convergence, and proposing error mitigation techniques to  overcome practical noise in NISQ devices.  of the two respective CNNs into a quantum fusion model,  which incorporates a QNN. Let us now briefly introduce the  individual components of the proposed architecture, which  is shown in Fig. 1.  *II-A1. 3D-CNN*  *Contributions:*  CNNs are deep learning models  specifically designed for processing and analyzing high  dimensional arrays such as images, volumes, or series data.  Because of their ability to automatically learn and extract  features from input data, they are a cornerstone of modern  computer vision applications. The adopted 3D-CNN model  in this work is based on the ResNet architecture with two  residual blocks [14]. After each layer, the output is passed  through a batch normalization followed by a nonlinear ReLU  activation function. The output from the convolutional layers  is then pooled, flattened, and fed through a sequence of fully  connected layers, producing the final output of the model.  Fig. 1(a) shows the exact architecture of the adopted 3D  CNN.  *II-A2. SG-CNN*  *Contributions:* An SG-CNN capitalizes on the benefits  of convolutional layers while leveraging the structural  relationships within protein-ligand complexes. In particular,  an SG-CNN effectively captures and preserves spatial  information using a 2D graph representation, where each  edge corresponds to a bond between atoms across all  molecules. For each molecule within the complex, spatial  information and associated features are initially processed  through a graph gated recurrent unit (GGRU), incorporating  information from its nearest neighbors. The resulting output  vector subsequently enters another GGRU, accumulating  information from the next nearest neighbors. This pivotal  stage, known as the *graph gather step*, is followed by the  data passing through a sequence of fully connected layers,  ultimately yielding the final output of the SG-CNN. The  exact architecture of the adopted SG-CNN is shown in  Fig. 1(b).  *II-A3. Quantum Fusion Model*  *Contributions:* The quantum fusion model, shown in  Fig. 1(c), takes the outputs from the second-to-last layer of  the two aforementioned CNN models. In particular, this input  is a vector of size 16, as it entails the output of 10 nodes from  the 3D-CNN and 6 nodes from the SG-CNN. By optimally  fusing these two pre-final layers, this strategy effectively  aggregates the acquired knowledge during the training of  the CNNs, thereby achieving superior performance. Unlike  the classical fusion model in Ref. [11], which is simply a  one-layer feed-forward neural network, the quantum fusion  model incorporates a QNN that consists of a quantum circuit  divided into two blocks. The first block is the *quantum*  *encoding* part, which maps the input data into a quantum  circuit, and the second block is a *parameterized quantum*  *circuit (PQC)*, where quantum operations are applied to  retrieve information from the encoded data. In the quantum  low expressibility. It outperforms classical counterparts  and some other PQC architectures, as shown in [22].  *•* **Circuit 2**. Constructed with a layer of Hadamard gates,  followed by all-to-all CZ connections and an additional  layer of rotations, this PQC boasts the highest entan  gling capacity among the circuits investigated in the  study from Ref. [20].  *•* **Circuit 3**. Composed by Y-axis rotations followed  by a layer of CNOT gates, this PQC exhibits high  entangling capacity and low expressibility. Frequently  categorized with *basic entangling layers*, it incorporates  fewer training parameters than Circuit 1.  *•* **Circuit 4**. Featuring two layers of Y-axis rotations  and two layers of controlled X-axis rotations, this  PQC achieves a significant balance between entangling  capacity and expressibility. However, it entails higher  complexity compared to Circuits 1-3.  *•* **Circuit 5**. Similar to Circuit 4, but incorporating Z-axis  rotations instead of X-axis rotations.  *•* **Circuit 6**. Comprising two layers of both X-axis and  Z-axis rotations, along with two layers of all-to-all  controlled rotations around the X-axis, this circuit was  the most expressive circuit in the study from Ref. [20].  Nevertheless, it also presents the highest complexity  among the considered circuits.  **II-B. Training Dataset**  The PDBbind dataset [23] (2020 version) is adopted as the  input to train the proposed quantum fusion model. PDBbind  represents an extensive compilation of experimentally deter  mined binding affinity data between proteins and ligands.  This dataset meticulously associates protein-ligand com  plexes with their respective affinity measurements, a curation  process executed via manual extraction from peer-reviewed  scientific publications. The latest PDBbind version, released  in early 2020, contains a total of 19,443 protein-ligand com  plexes. Additionally, a meticulously curated subset of 5,316  samples has been compiled, specifically comprising high  quality complexes. Finally, an even higher-quality core set  of 285 samples is derived, primarily for validating binding  affinity prediction methods. We utilize the refined set in the  training and validation phases of our analysis, with 25% of  the data reserved for validation. Then, the core set is used  for the testing phase.  **II-C. Data Pre-processing**  Before passing the PDBbind data into the two CNN  architectures, it is essential to pre-process the raw data and  extract pertinent input features from the 3D structures of  proteins and ligands. In this regard, our approach closely  aligns with established featurization techniques as outlined  in [11], [13], [24], thereby facilitating the comparability of  our results with existing state-of-the-art models. Moreover, | In the field of drug discovery, it is imperative to identify  proteins that are instrumental in the cascade of molecular  interactions leading to a specific disease [2]. Upon the iden  tification of such a target protein, a list of prospective drug  candidates is generated. These candidates, often described  as small molecules or compounds termed *ligands*, have the  potential to modulate the target protein’s activity through  binding interactions [3]. Ideal ligands are chosen based on  their high binding affinity to the target protein, coupled with  minimal off-target interactions with other proteins. However,  quantifying such binding affinities is a resource-intensive  endeavor [4], both in terms of time and financial investment.  This is particularly true considering that the initial screening  process often encompasses thousands of compounds [4], [5].  The transition from conventional laboratory methods to  computer-aided design (CAD) has markedly improved the  efficiency and accuracy of drug discovery and binding  affinity prediction. This advancement has been further bol  stered by the incorporation of artificial intelligence (AI)  and machine learning (ML) algorithms, which facilitate  exhaustive analyses of large-scale datasets, uncovering pre  viously undetected patterns related to the atomic features  of protein-ligand molecular complexes [6]. Furthermore,  recent strides in quantum computing have added another  layer of sophistication to drug discovery efforts, offering  unprecedented parallelized computational capabilities [7].  Quantum machine learning (QML) models, in particular,  are well-suited to manage the challenges of exponentially  increasing data dimensionality, often outperforming tradi  tional ML models under specific conditions. Taken together,  these technological advancements make QML and hybrid  quantum-classical models highly promising for navigating  the complex, high-dimensional challenges intrinsic to drug  discovery [8]. The main challenge facing such QML mod  els is the noise prevalent in today’s *noisy intermediate*  *scale quantum (NISQ)* computers, which requires advanced  quantum error mitigation techniques to yield a scalable and  accurate performance.  *Related Works:* Several prior works [9]–[13] addressed  the problem of binding affinity prediction in drug discovery  using tools from both classical ML and QML. For instance,  *Contributions:* The main contribution of this paper is  the development of a novel hybrid quantum fusion model  aimed at enhancing the binding affinity prediction in drug  discovery. The proposed model strategically integrates 3D  CNNs and SG-CNNs to leverage their respective strengths in  processing diverse facets of the training data. The proposed  quantum architecture is meticulously designed for optimal  accuracy. Simulation results demonstrate the superior per  formance of the proposed hybrid quantum fusion model  relative to state-of-the-art classical models. Particularly, the  proposed model achieves a 6% improvement in the binding  affinity prediction accuracy, and exhibits faster, smoother,  and more stable convergence, thereby boosting its general  ization capacity. To enhance the scalability of the proposed  quantum fusion model on NISQ devices, we present a novel  error mitigation technique capable of effectively alleviating  noise introduced in quantum circuits for error probabilities  *p ≤* 0*.*05. This method incurs no additional overhead during  both the training and testing phases, paving the way for  broader QML applications over NISQ devices.  **II. SYSTEM MODEL**  This section describes the proposed hybrid quantum fusion  model, its components, the data used in training the model,  and the necessary pre-processing steps.  **II-A. Proposed Hybrid QML Architecture**  The proposed hybrid quantum fusion architecture builds  upon the classical fusion model introduced in [11] while  integrating quantum neural networks (QNNs) into the model  design. In particular, the protein-ligand complex data (see  Section II-B) is initially fed into a 3D-CNN and an SG  CNN, simultaneously. Then, late-late fusion is performed by  feeding the outputs from the second-to-last layer of each  fusion model, the considered outputs of the 3D-CNN and  SG-CNN models are first embedded into quantum states  through quantum encoding circuits, then run through the  trainable PQC.  *Quantum Encoding Techniques:* A variety of quantum  data encoding techniques have been recently developed in  the literature [15]. In this work, we focus on two of the most  effective encoding techniques, analyzing and comparing their  performance in the context of binding affinity prediction.  Future works will consider alternative dimensionality re  duction techniques, like the tensor train network [16], for  more general higher-dimensional data to guarantee optimal  training. The two considered methods are:  1) **Amplitude encoding**, where the features are en  coded in the amplitudes of the quantum state in the  computational basis [17]. This scheme requires only  *⌈*log2 (*n*)*⌉* qubits to encode a data sample into a  quantum state, where *n* represents the input dimension  of the QNN. The depth of the embedding circuit grows  as *O*(*poly*(*n*)) while the number of parameters subject  to optimization scales as *O*(*log*2(*n*)).  2) **Hybrid Angle Encoding (HAE)**, where amplitude  encoding is implemented using parallel blocks of in  dependent qubits [15]. The features are divided into  *b* blocks of size 2 *m −* 1, where *m* is the number  of qubits. Accordingly, *b × m* qubits are required to  encode the whole data sample into a quantum state  using the HAE.  *PQC Architecture:* After encoding the data into quan  tum states (or qubits), they pass through a PQC. The design  of the PQC is crucial to guarantee optimal performance of  the quantum fusion model [18], [19]. The PQC consists of  several quantum gates which are controlled by classically  optimized parameters. A notable challenge is to choose  an effective circuit that adequately represents the solution  space while minimizing circuit depth and the number of  parameters. Up to this point, two robust metrics have been  introduced to assess the quality of PQCs [20], as discussed  next.  The first metric, termed *expressibility*, gauges the PQC’s  capacity to explore the Hilbert space, thereby generating a  diverse array of quantum states [20]. The second metric,  termed *entangling capacity*, quantifies the PQC’s ability to  generate entangled states [21].  In this work, we examine six distinct PQC architectures,  each characterized by differing levels of expressibility, entan  gling capacity, and number of training parameters. The goal  is to identify the most optimal architecture for our quantum  fusion model. Each PQC is composed of *L* layers of quantum  gates, which are depicted in Fig. 2.  *•* **Circuit 1**. Comprising 3D rotation gates with all-to-all  CNOT connections forming *strongly entangling layers*,  this circuit demonstrates high entangling capacity and  the PDB files are converted to Mol2 files using the UCSF  Chimera sofware [25].  The input data for 3D-CNN consists of spatial represen  tation of 3D structures of atoms of protein-ligand pairs.  The atoms are voxelized into a grid of size *N × N × N*  with a voxel size of 1  ˚  A, and the number of voxels set to  *N* = 48 to strike a balance between covering the entire  pocket regions and lowering the input data size for the CNN  model. *C* = 19 features are extracted for each voxelized  atom using OpenBabel [26], bringing the input data for 3D  CNN to a *C × N × N × N* matrix. Such features include  the atom type, hybridization, number of heavy atom bonds,  number of bonds with heteroatoms, structural properties,  partial charge, and molecule type (protein vs ligand). On the  other hand, the input data for SG-CNN consists of a spatial  graph representation of the protein-ligand complexes, with  the atoms forming the nodes, and the bonds between atoms  forming the edges, of the graph. As in Ref. [11], covalent  bonds are represented by an *N × N* adjacency matrix and  non-convalent bonds by an *N ×M* matrix, where the element  *Aij* of each matrix represents the Euclidean distance between  the atoms *i* and *j*. |