**Accelerating Drug Discovery through Deep Learning and AI: A Mini Review**

**Abstract:**

Artificial intelligence and deep learning models have recently demonstrated vast capabilities to enhance and accelerate drug discovery across areas including predicting molecular interactions, generating optimized compounds, repurposing agents, and analyzing complex biomedical data. This mini review synthesizes major concepts, approaches, applications and findings on the role of AI/deep learning in drug discovery highlighted across a collection of 85 papers from 2019-2023. Key themes center on target prediction, de novo design, pharmacokinetics, repurposing, and multimodal data integration. We highlight marked improvements demonstrated over previous computational methods and experimental pipelines. Challenges still being addressed encompass data limitations, evaluation metrics, model interpretability, and translational validation. The level of innovation already achieved points to AI/deep learning techniques becoming indispensable digital tools to increase efficiency and productivity in therapeutic development. Ongoing areas of promise include multimodal analysis to contextualize chemical-biological interactions, molecular generative model optimization, and real-world evidence applications.

**Keywords:**

**Deep learning, Artificial intelligence, Drug discovery, Drug design, Drug development, Drug repurposing, Molecular modeling, Computational chemistry, Systems biology, Multimodal data**

**Introduction**

Artificial intelligence (AI) and deep learning have emerged as powerful and promising techniques to accelerate and enhance many aspects of drug discovery and development. Over the past decade, a proliferation of studies have demonstrated the capabilities of AI/deep learning in improving target identification, hit discovery, lead optimization, and the prediction of pharmacokinetic properties and toxicity (Born & Manica, 2021; Gupta et al., 2021). This mini-review synthesizes major concepts, approaches, applications and findings on the role of AI/deep learning in drug discovery highlighted across the provided collection of 85 papers spanning 2019-2023.

At a high level, the key themes and capabilities of AI/deep learning center around: 1) Predicting drug-target interactions, bioactivity and affinity; 2) Generating novel molecular structures and optimizing compounds; 3) Predicting pharmacokinetic properties; 4) Drug repurposing; and 5) Integrating and analyzing complex, multi-modal biomedical data. Across these areas and specific applications, studies consistently showcase marked improvements over previous computational methods and the ability to accelerate discovery along both the computational and experimental pipelines.

**Predicting Drug-Target Interactions, Bioactivity & Affinity**

A predominant application of AI/deep learning involves predicting interactions and quantifying the affinity or activity between drug compounds and protein targets, with implications for hit identification, lead prioritization and optimization, and repurposing (Chen et al., 2021; Fu et al., 2021; Gupta et al., 2021; Joshi et al., 2021; Monteiro et al., 2021; Wang et al., 2021). For example, Chen et al. (2021) combined graph neural networks and recurrent neural networks in an end-to-end architecture that improved AUC by 2.4% and recall by 9.4% over previous methods in predicting drug-target interactions on imbalanced datasets. Joshi et al. (2021) proposed a generative model utilizing 3D target scaffold information that produced compounds with more favorable predicted binding vs. decoys. Overall, a comprehensive assessment of drug-target interaction prediction methods underscores deep learning as state-of-the-art (Zuo et al., 2021). Ongoing challenges that emerging studies aim to address include lack of quality labeled data, hidden bias, model interpretability, and transferability to new data (Wang et al., 2022; Zhao et al., 2023).

**De Novo Molecular Generation & Optimization**

Deep generative models including variational autoencoders (VAEs), generative adversarial networks (GANs), and reinforcement learning (RL) strategies have attractions for directly producing optimized, novel molecular structures rather than screening existing libraries (Bian & Xie, 2021; Born & Manica 2021). Though relatively early in development and adoption, innovative applications highlight the promise. For example, Zhong et al. (2021) achieved 95% validity in generating GPCR-focused structures using a VAE-GAN architecture. DL-based RL optimization further improved docking scores by 10-20% over compounds of similar size from traditional methods (Xiong et al., 2023). Key remaining challenges encompass data scarcity limitations, evaluation metrics, interpretability, and experimental validation (Wang et al., 2022).

**Prediction of Pharmacokinetic & Toxicity Properties**

Multiple studies demonstrate AI/deep learning methods accurately predicting key pharmacokinetic endpoints such as clearance, bioavailability, and plasma protein binding (Huang et al., 2021; Iwata et al., 2021; Khaouane et al., 2023), with specific architectures tailored to learn from the structure-activity relationships. For toxicity evaluation, Li et al. (2021) achieved top performance in predicting drug-induced liver injury (DILI), with a deep framework combining model-based descriptors. Deep learning analysis applied to cellular images further differentiated mechanisms of compound-induced cytotoxicity (Metzger et al., 2022).

**Drug Repurposing**

A major motivation for computational drug repositioning includes shortcutting lengthy experimental timelines by identifying new uses for approved or investigational drugs. For example, Liu et al. (2021) designed a deep learning pipeline emulating clinical trials for coronary disease that revealed drugs and combinations offering substantial outcome improvements without previous indications. Multiple studies particularly focus on repurposing for cancers (Issa et al., 2021; Magge et al., 2021) and COVID-19 treatment via predicted inhibition of SARS-CoV-2 proteins (Azmoodeh et al., 2022; Jin et al., 2021; K et al., 2021).

**Integrating & Analyzing Multimodal Data**

A consistent theme throughout is deep learning’s capability to integrate and model relationships within large, diverse, multimodal biomedical data to reveal insights for drug discovery. For instance, Allesoe et al. (2023) combined clinical markers, genetics, proteins, metabolites and gut microbiome profiles to discover drug-omic associations and opposite molecular responses for related statins. Multitask networks and autoencoders further boost performance by jointly analyzing cell line genomic data for predicting chemical synergy and bioactivity (Schwarz et al., 2021; Zhang et al., 2021).

**Conclusion & Future Outlook**

In summary, deep learning/AI is demonstrating vast and still expanding potential to transform computational approaches, inform experimentation, and accelerate pipelines across the spectrum of drug discovery objectives. Multimodal analysis to contextualize chemical-biological interactions, optimization of molecular generative models, and translational applications leveraging real-world evidence are compelling areas warranting additional investigation. Integrating advanced AI and systems biology with closed-loop experimental validation will contribute to a new era of enhanced efficiency, productivity and insight in therapeutic development. While still early in adoption, the level of innovation and progress achieved thus far signifies these techniques likely to become indispensable digital tools in the drug discovery arsenal.