IECC 2026 Conference
The 5th International Electronic Conference on Cancers
10–12 June 2026 | Online
MDPI

NDS: A Novel Deep Learning-Based Systems Biology Framework for Identifying Prognostic Biomarkers in Hepatocellular Carcinoma

Muhammad Zurgham Akram[1], Mehwish Majeed[1]

[1]College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, Liaoning, 110819, China.

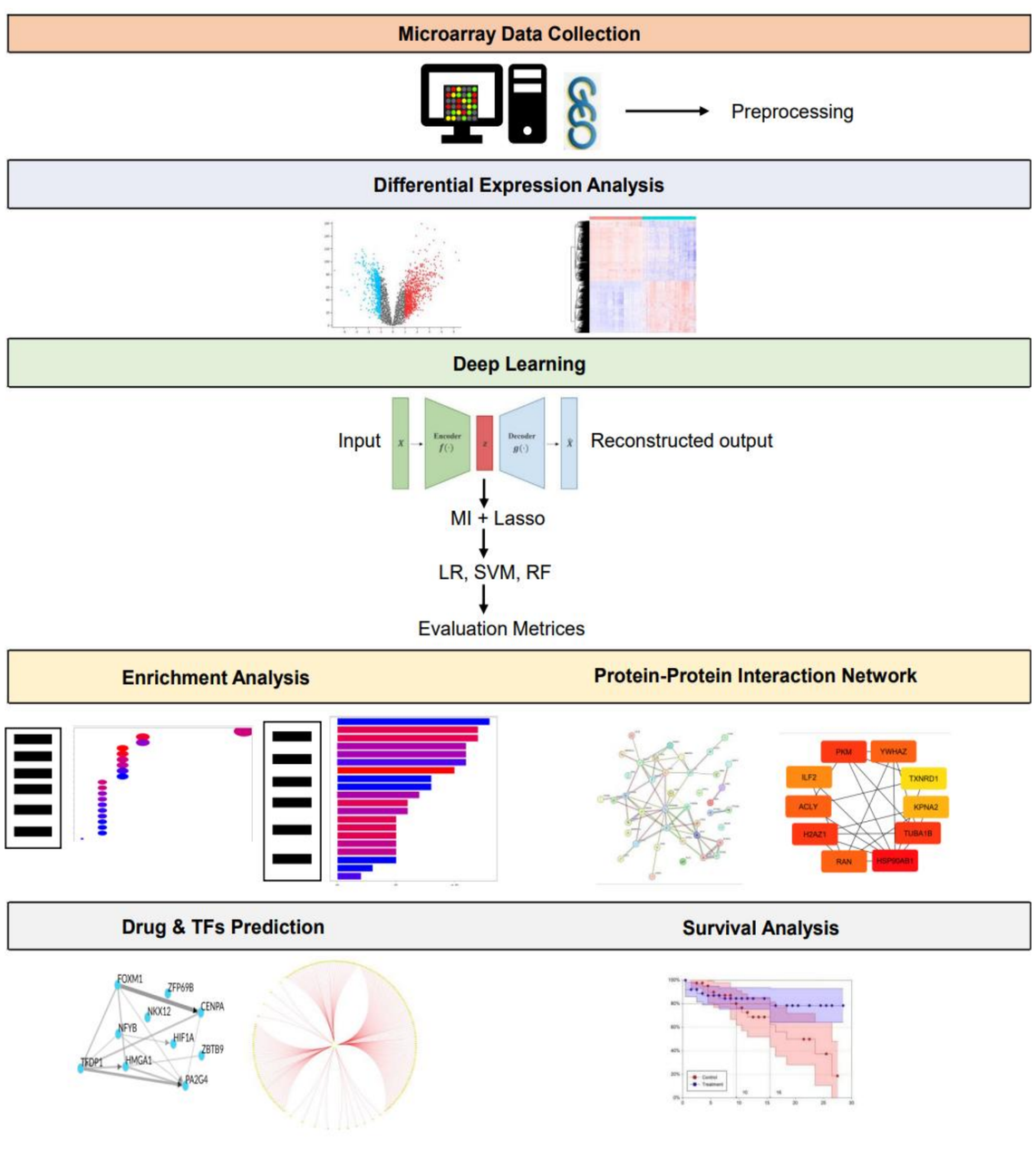
INTRODUCTION & AIM

Hepatocellular Carcinoma (HCC) is one of the most aggressive and prevalent forms of liver cancer and accounts for nearly 75–85% of all liver cancer cases worldwide. Despite advances in surgical resection, liver transplantation, and targeted therapies, the prognosis of HCC patients remains poor because most cases are diagnosed at advanced stages. Therefore, the identification of reliable molecular biomarkers is essential for improving early diagnosis, prognosis, and precision treatment strategies.

Traditional bioinformatics and machine learning approaches often struggle with high-dimensional transcriptomic data and may fail to capture complex nonlinear interactions among genes. Recent advances in deep learning, particularly autoencoder-based dimensionality reduction, provide powerful tools for extracting biologically meaningful latent features from large-scale gene expression datasets.

The present study aimed to develop an integrative deep and machine learning framework for the identification of reliable prognostic biomarkers in HCC by analyzing differential gene expression profiles and extracting highly informative latent features associated with tumor progression and patient survival.

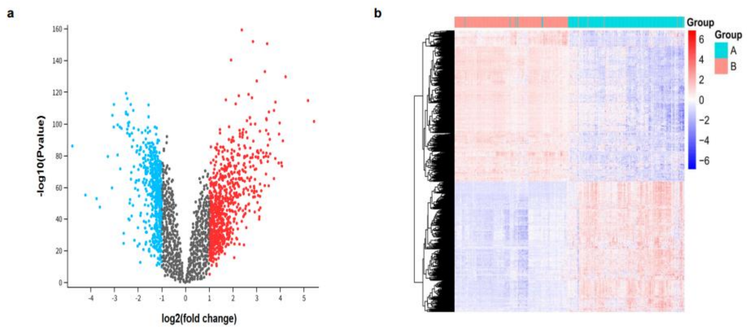
METHODOLOGY



RESULTS & DISCUSSION

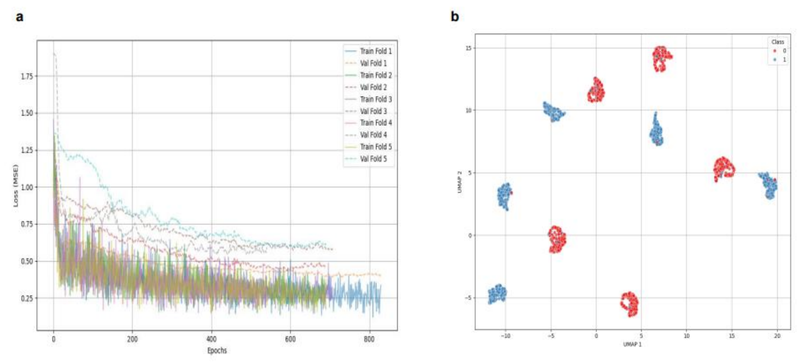
Differential Expression Analysis

A total of 1,380 DEGs were identified between HCC and normal tissues, demonstrating substantial transcriptomic alterations associated with tumor development.



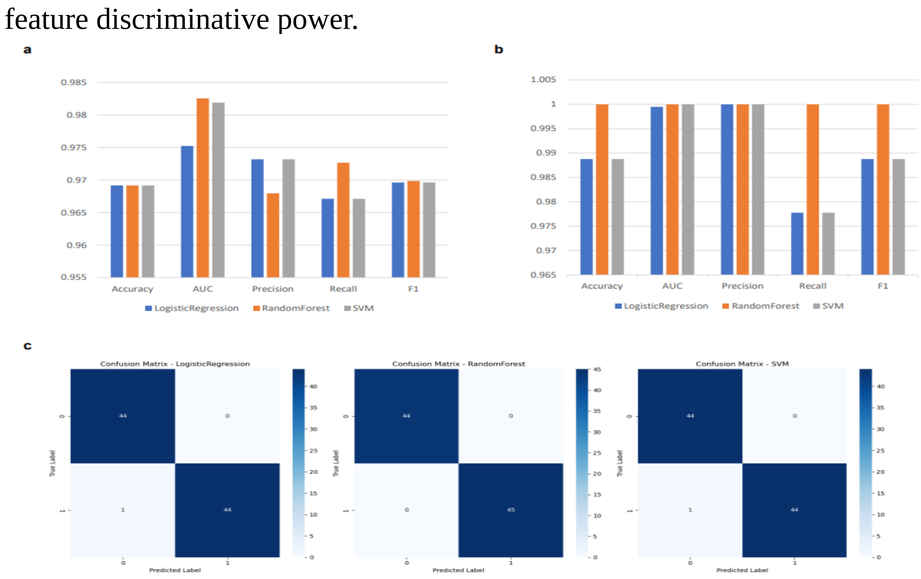
Autoencoder and Feature Selection Performance

The deep autoencoder effectively reduced data dimensionality while preserving biologically relevant information. UMAP visualization demonstrated clear separation between tumor and normal samples within the latent feature space. The MI-LASSO pipeline successfully refined predictive features and selected 50 highly informative genes for downstream analysis.



Classification Performance

Cross-validation showed stable and high performance across classifiers (accuracy = 0.969, F1 ≈ 0.97), with RF and SVM achieving slightly higher AUC (~0.983 and ~0.982). Independent tests showed excellent performance across all classifiers (~1.0), confirming high feature discriminative power.



Functional Enrichment Analysis

Enrichment analysis revealed involvement in metabolic pathways, PI3K/Akt signaling, protein processing in the endoplasmic reticulum, fatty acid metabolism, cell cycle regulation, and viral carcinogenesis.

Identification of Hub Genes

PPI network analysis identified ten key hub genes: *HSP90AB1*,*TUBA1B*, *PKM*, *H2AZ1*, *YWHAZ*, *ACLY*, *RAN*, *ILF2*, *KPNA2*, *TXNRD1*.These genes exhibited strong interaction connectivity and significant prognostic relevance.

Survival and Therapeutic Significance

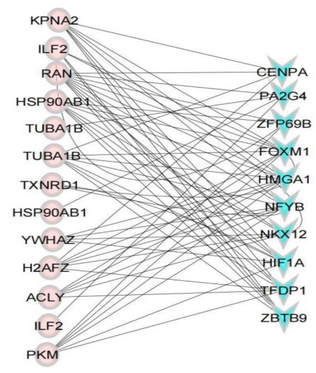
Survival analysis showed that hub genes, were significantly associated with poor prognosis (HR > 1.5, p < 0.05), correlating with reduced overall, relapse-free, and disease-specific survival in HCC.

TFs analysis

ChEA3 tool identified key regulators of the ten hub gene

Core Gene predictive drug Analysis

Through DGIdb database the drugs screening reveals *TUBA1B* associates with 83 drugs, *HSP90AB1* with 46 drugs, *PKM* with 36, *TXNRD1* with 6, and *H2AZ1* and *ACLY* with 4,4 drugs, respectively.



CONCLUSIONS

This integrative novel framework effectively identifies biologically relevant biomarkers, providing insights into HCC mechanisms and potential targets for precision therapy.

FUTURE WORK

Future works should focus on experimental validation of identified biomarkers using in vitro and in vivo studies.