Journal of Universal Computer Science, vol. 31, no. 2 (2025), 159-188
submitted: 11/4/2024, accepted: 27/8/2024, appeared: 28/2/2025 CC BY 4.0

Autoencoder-Integrated WideResNet with Dynamic Optimization (AIW-DynOpt): A Novel Hybrid Deep Learning Approach for Head and Neck Cancer Gene Expression Analysis

Aneela Nargis
(DHA Suffa University, Karachi, Pakistan
aneela.nargis@live.com)

Muhammad Mobeen Movania
(Habib University, Karachi, Pakistan
mmmovania@gmail.com)

Shama Siddiqui
(DHA Suffa University, Karachi, Pakistan
shamasid@hotmail.com)

Abstract: Head and neck cancer presents a significant global health challenge, necessitating the development of robust computational models for accurate gene expression analysis. This study introduces Autoencoder-Integrated WideResNet with Dynamic Optimization (AIW-DynOpt), a novel hybrid deep learning framework specifically designed for analyzing head and neck cancer gene expression data. AIW-DynOpt integrates a Deep Undercomplete Autoencoder (DUAE) with a Wide Residual Network (WideResNet) architecture and employs the Successive Halving algorithm for optimal model selection. Utilizing the Cancer Genome Atlas (TCGA) HNSC dataset, which comprises 20,503 genes and 564 samples, our approach focuses on enhancing predictive performance and computational efficiency. A comprehensive evaluation of AIW-DynOpt was conducted, benchmarking it against alternative methods such as DUAE paired with Convolutional Neural Networks (CNN), Support Vector Machines (SVM), and Naive Bayes classifiers. The results demonstrate that AIW-DynOpt consistently outperforms the alternative methods across multiple performance metrics, including accuracy, recall, sensitivity, specificity, and Area under the Curve (AUC). Additionally, AIW-DynOpt exhibits superior computational efficiency, significantly reducing model training time while maintaining high predictive accuracy. This study underscores the potential of hybrid deep learning frameworks in advancing computational models for cancer research, positioning AIW-DynOpt as a promising tool for precise gene expression analysis in head and neck cancer and beyond.

Keywords: Gene profiling, Bioinformatics, Deep fusion, Hyperparameter optimization, Performance accuracy
Categories: H.3.1, H.3.2, H.3.3, H.3.7, H.5.1
DOI: 10.3897/jucs.125224

1 Introduction

Head and neck cancer (HNC) constitutes a diverse group of malignancies affecting various structures in the head and neck region, including the oral cavity, pharynx,

larynx, and paranasal sinuses. The primary risk factors for HNSCC include tobacco use, alcohol consumption, and infection with high-risk types of human papillomavirus (HPV), particularly HPV-16 [Bakhtiar, 15; Beck, 16]. The prevalence of HPV in oropharyngeal squamous cell carcinoma (OPSCC) ranges from 50% to 90%, significantly influencing the disease's molecular characteristics and clinical outcomes [Bakhtiar, 15; Beck, 16; Stephen, 15].

HNSCC is characterized by a high degree of genomic instability and a substantial burden of mutations, contributing to the heterogeneity observed within and between tumors. Large-scale genomic studies, such as those conducted by The Cancer Genome Atlas (TCGA), have identified frequent mutations in key oncogenes and tumor suppressor genes. Notable mutations include TP53, which is frequently mutated in HNSCC and encodes a critical tumor suppressor involved in DNA repair, cell cycle regulation, and apoptosis [Poeta, 07]. Mutations in TP53 lead to the loss of its tumor suppressor functions, contributing to genomic instability and cancer progression. CDKN2A, another commonly mutated gene, encodes p16INK4a, a protein that inhibits cyclin-dependent kinase 4 (CDK4) and regulates the cell cycle. Mutations or deletions in CDKN2A disrupt cell cycle control, facilitating uncontrolled cell proliferation [Reimers, 07]. Mutations in PIK3CA activate the PI3K/AKT pathway, promoting cell growth and survival, and are prevalent in a subset of HNSCC cases, representing potential targets for therapeutic intervention [Leemans, 18]. Additionally, mutations in NOTCH1 can either activate or inactivate the NOTCH signaling pathway, influencing cell differentiation and proliferation, underscoring the complexity of the disease [Agrawal, 11; Ju, 23; Tapak, 23].

The rising incidence of HPV-associated HNSCC has introduced a new dimension to the molecular landscape of head and neck cancer. HPV-positive (HPV+) HNSCC represents a distinct subtype with unique molecular features compared to HPV-negative (HPV-) HNSCC. The viral oncogenes E6 and E7, expressed in HPV+ tumors, inhibit tumor suppressors such as TP53 and RB1, contributing to carcinogenesis [Fakhry, 08; Tapak, 23]. HPV status has significant prognostic implications, with HPV+ HNSCC patients exhibiting better responses to chemoradiotherapy and a lower risk of death compared to HPV- patients. This difference in outcomes is partly due to the distinct biological behavior of HPV+ tumors [Settle, 09].

There are notable racial disparities in the incidence and outcomes of HNSCC. African Americans (AA) exhibit higher incidence and mortality rates compared to Caucasian Americans (CA) [Settle, 09; Bakhtiar, 15; Beck, 16; Stephen, 15]. Studies indicate that the poorer survival outcomes in AA with oropharyngeal tumors are primarily due to a lower prevalence of HPV+ tumors in this population. When stratified by HPV status, AA and CA patients with HPV- tumors show similar survival outcomes [Settle, 09; Bakhtiar, 15; Beck, 16; Stephen, 15].

Epigenetic alterations, such as DNA methylation, histone modifications, and non-coding RNA regulation, play a significant role in HNSCC. Promoter hypermethylation of tumor suppressor genes is an early event in tumorigenesis, serving as a potential biomarker for cancer detection. HPV+ HNSCC tumors exhibit distinct epigenetic profiles, including genome-wide hypomethylation and promoter hypermethylation [Ha & Califano, 16; Stransky, 11; Tapak, 23]. Recent studies have applied pathway analysis to investigate the biological roles of differentially methylated genes in HPV+ and HPV- HNSCC, identifying key signal transduction pathways relevant to the disease [Worsham, 13; Bakhtiar, 15; Beck, 16; Stephen, 15]. It represents a significant global

health burden, with approximately 890,000 new cases diagnosed worldwide each year [Bray, 18]. Despite advancements in treatment modalities, the prognosis for HNSCC patients remains variable, underscoring the need for precise diagnostic and prognostic tools [Leemans, 18].

Gene expression analysis plays a pivotal role in understanding the molecular mechanisms underlying HNSCC progression and identifying potential therapeutic targets. High-throughput technologies, such as microarray and RNA sequencing, enable the simultaneous measurement of expression levels for thousands of genes across HNSCC samples [Brierley, 17]. However, the complexity and dimensionality of gene expression data pose significant challenges for traditional analytical approaches [Lee, 23; Thudumu, 20; Fan, 14; Rahnenführer, 23]. This complexity is compounded by the critical need to discern subtle, yet biologically significant patterns that are often pivotal for diagnostic, prognostic, and therapeutic decision-making [Naylor, 10; Sebastian, 91; Garcia, 23].

The Cancer Genome Atlas (TCGA) HNSCC dataset represents a valuable resource for studying the molecular signatures of HNSCC [Lawrence, 15]. Comprising gene expression data from thousands of samples, TCGA offers a comprehensive view of the genomic alterations associated with HNSCC development and progression. Using such large-scale datasets can yield insights into the biological underpinnings of HNSCC and facilitate the development of novel therapeutic interventions. The results shown in this research are based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.`

Consequently, there is a growing interest in involving advanced computational techniques, particularly deep learning, to extract meaningful insights from HNSCC gene expression data. Deep learning models, characterized by their ability to automatically learn hierarchical representations from raw data, have demonstrated remarkable success in various biomedical applications, including cancer diagnosis and prognosis [Sarker, 21; Ahmed, 23; Spratling, 23; Esteva, 17]. In addition to traditional deep learning architectures, researchers have explored the realm of hybrid models, combining different methodologies to address complex data analysis tasks. One notable example is the integration of dimension reduction techniques with classification algorithms, as demonstrated by [Kilicarslan, 20; Liu, 17; Saha, 22], aiming to streamline the analysis of microarray data in cancer research. This hybrid approach effectively reduces the dimensionality of the microarray dataset, thereby enhancing the efficiency of feature selection and bolstering the overall performance of subsequent Convolutional Neural Network (CNN) classifiers.

Motivated by the need for robust computational models to analyze HNSCC gene expression data, our research aims to develop a novel hybrid deep learning framework Autoencoder-Integrated WideResNet with Dynamic Optimization (AIW-DynOpt). By integrating multiple deep learning components, including deep undercomplete autoencoder and WideResNet, with optimization algorithms, such as Successive Halving, we seek to enhance the accuracy and interpretability of predictive models for HNSCC diagnosis and prognosis. Current diagnostic approaches often face limitations in effectively utilizing genomic data to differentiate between tumor types within HNSCC. AIW-DynOpt is designed to overcome these limitations by employing advanced deep learning techniques that can capture and utilize intricate patterns inherent in gene expression data.

Compared to traditional methods like SVM, CNN, and Naive Bayes classifiers, AIW-DynOpt offers several advantages. It consistently demonstrates superior performance metrics, including accuracy, recall, sensitivity, specificity, and Area under the Curve (AUC). Moreover, AIW-DynOpt reduces training time without compromising predictive accuracy, making it suitable for analyzing large-scale genomic datasets. To the best of our knowledge no such models have been developed in the past. The major contributions of the present research are listed below:

- Introducing the AIW-DynOpt hybrid model, which combines the power of Autoencoder-Integrated WideResNet with Dynamic Optimization, presenting a novel approach to gene expression analysis in head and neck cancer research.
- Demonstrating superior classification performance compared to traditional methods such as SVM and CNN, as evidenced by higher accuracy, precision, recall, F1 score, and AUC score.
- Utilizing the Successive Halving algorithm for hyperparameter optimization, which efficiently explores the hyperparameter search space, leading to improved model performance without excessive computational costs.

In the subsequent sections of this paper, Section II presents the review of existing literature; Section III describes the methodology used to develop AIW-OptDyn model; Section IV presents experimental results and a comparative analysis with existing approaches; Section V discusses the implications of our findings for HNSCC research and clinical practice; and finally, Section VI concludes the paper and offer directions into future research.

2 Related Work

In the domain of head and neck cancer classification, various machine learning and deep learning models have been explored to enhance diagnostic accuracy and prognostic predictions. Firstly, the predictive models for disease status and early-stage detection of head and neck squamous cell carcinoma (HNSCC) outlined by [Kaddi, 15], demonstrated effectiveness, yet limitations such as data source and computational complexity were noted. Meanwhile, the Variational Autoencoder (VAE) model used by [Way, 18], for cancer gene expression data captured biologically relevant features but required rigorous experimental testing and validation. Furthermore, the study by [Lyu, 18], presents a convolutional neural network (CNN) model for classifying 33 tumor types including head and neck cancer using gene expression data from the Pan-Cancer Atlas. While the CNN model achieved high accuracy, challenges such as high dimensionality, complexity for multiple classes, interpretability, and generalization were noted. In another investigation by [Patil, 19], machine learning techniques were applied to genomic data in head and neck cancer (HNC) for prognostic prediction, such as support vector machine (SVM), adaptive neuro-fuzzy inference system (ANFIS), and artificial neural networks (ANN). Despite achieving promising accuracy rates, challenges such as limited data samples and lack of standardized study designs hindered the integration of these models into clinical practice for HNC prognostic prediction.

[Carenzo, 20], defines a 13-gene model focused on predicting the survival of HPV-negative oral cavity squamous cell carcinoma (OSCC) patients. This model, initially developed by [Lohavanichbutr, 09] and revised by Chen et al. in 2009, incorporates clinical features such as age and gender along with the 13-gene signature. The model

significantly stratified patients into higher and lower risk groups, demonstrating its potential for improving prognostic accuracy. Moreover, Zhao et al. in 2020, presented the deep learning-based prognostic model integrating multi-omics data using an autoencoder, showcased robustness in predicting disease progression in HNSCC patients, although challenges in data standardization and interpretation were noted. Similarly, convolutional neural network (CNN) models have been employed to classify tumor samples and identify cancer markers, demonstrating high prediction accuracies [Mostavi, 20]. The model defined by Xie et al. in 2020, is a multifactorial deep learning model. The model integrates tumor mutation burden, microsatellite instability, and somatic copy-number alterations to classify tumors of different types into different genomic clusters. The model utilizes deep learning techniques, specifically Deep Belief Networks (DBN) and Denoising Autoencoders (DAE), to extract features from multifactorial inputs and perform stratification analysis to identify distinct genomic clusters.

Furthermore, the deep neural network (DNN) classifier coupled with CIBERSORT, as described by Kim et al. in 2021, demonstrated efficacy in predicting survival patterns based on tumor-infiltrating lymphocyte (TIL) fraction patterns derived from RNA expression data of oral cancer patients. While promising, concerns regarding model applicability to other cancer types and dataset quality were noted. Meanwhile, the application of deep learning algorithms, particularly convolutional neural networks (CNN), in multi-omics data analysis for head and neck tumor diagnosis, as discussed by [Wang, 21], highlighted the potential of transfer learning and the need for further research in integrating diverse omics data. Additionally, the CNN model was developed by [Darendeli, 21], utilizing gene expression data transformed into RGB format, exhibited high accuracy in cancer prediction, yet concerns regarding dataset diversity and generalizability remained.

In the systematic review reported by [Alabi, 21] on machine learning in oral squamous cell carcinoma, various machine learning models are applied to gene expression datasets to aid in diagnosis, prognosis, and prediction tasks. Some of the commonly used algorithms include: Support Vector Machine (SVM), Artificial Neural Network (ANN), and Ensemble Methods. These models extract valuable insights from gene expression datasets, yet face challenges such as integration into clinical practice, explainability, data quality, algorithmic bias, and external validation. [Pratama, 21] developed a CNN model for analyzing SCCs based on gene expression showed limitations in differentiating oral squamous cell carcinoma (OSCC) from non-oral head and neck squamous cell carcinoma (HNSCC), indicating challenges in distinguishing between these two cancer types. The ResNet50 model applied by [Oza, 21], for predicting head and neck cancer demonstrated high performance, yet challenges related to data standardization and interpretation were acknowledged. In the realm of molecular subtyping, deep learning frameworks have enabled the identification of subtypes of oral squamous cell carcinoma (OSCC) based on gene expression profiles [Li, 21]. These subtypes exhibit significant differences in survival outcomes, offering insights for personalized treatment strategies.

Moreover, the utilization of various machine learning models, including Support Vector Machine, Naïve Bayes, Decision trees, Perceptron, Logistic Regression, and K Means, Radial Basis Function Support Vector Machine (RBF-SVM), Random Forest, k-Nearest Neighbors (kNN), as outlined by [Kim, 21] and [Singh, 22], for head and neck and specifically Gingivobuccal Cancer (GBC) prognosis by [Singh, 22], indicated

promising avenues for predictive modelling, albeit with challenges such as lack of well-annotated attributes and data scarcity. Moreover, the integration of autoencoder and capsule network models described by [Wu, 22] for identifying gene signatures in head and neck cancer from multi-omics data emphasized the potential of deep learning techniques in elucidating molecular mechanisms, while acknowledging the need for further validation and optimization. [Sharma, 22] presented an integrative model for improved prediction accuracy of clinical outcomes in the survivability of oral cancer showcased promising predictive performance but faced limitations in data types, interactions, and dataset source. The artificial neural network (ANN) model presented by [Gupta, 22] for cancer prediction exhibited effectiveness but encountered challenges related to data size constraints, complexity of gene expression data, optimizer sensitivity, generalization to new datasets, and interpretability of results. A promising approach involved combining downsampling local binary patterns (DS-LBP) with Long Short-Term Memory (LSTM) networks for classifying heart conditions, achieving high classification rates but emphasizing the need for parameter optimization . Additionally, computational methods for HNSCC have screened natural compounds to alter gene expressions related to tumor regression, highlighting their potential in reducing drug resistance and enhancing immunotherapeutic efficacy [Kaya, 22 a, Akdağ, 22]. Further advancements include a hybrid model combining Factor Analysis (FA) and Extreme Learning Machine (ELM) for diagnosing various diseases, demonstrating improved diagnostic accuracy by reducing dataset dimensionality while maintaining high success rates [Kaya, 22 b]. the recent COVID-19 pandemic has seen the application of artificial intelligence methods for disease detection. A novel approach using x-ray images, the angle transform (AT) method, and a hybrid deep learning model combining GoogleNet and LSTM achieved a high classification accuracy of 98.97% for COVID-19 detection. This method, tested on a dataset from the Mendeley database, demonstrates the potential of deep learning and novel transformation techniques in medical image analysis [Kaya, 22 c].

In contrast, the artificial neural network model, neuralHNSCC, developed by [Luo, 23] showcased exceptional performance in classifying HNSCC using combined gene expression datasets. However, limitations such as small sample size, data source bias, incomplete clinical information, tissue source variation, and the need for further validation with diverse datasets were acknowledged. Similarly, the random forest model by [Weusthof, 2023] identified a 44-gene set related to perineural invasion (PNI) in HNSCC, aiding in risk assessment. Yet, retrospective study design and reliance on genome-wide transcriptome data raised questions about generalizability and clinical implementation. Similarly, the autoencoder model employed by [Tapak, 23], for feature extraction from gene expression profiles related to oral squamous cell carcinoma (OSCC) underscored the need for addressing overfitting and enhancing feature selection methods. Machine learning techniques have also found increasing utility in oral cancer diagnosis, as evidenced in a study given by [Dixit, 23], exploring various methods such as Logistic Regression, Decision Trees, and Support Vector Machines. These models enable the analysis of complex data patterns from imaging scans or genetic markers, facilitating early cancer detection and personalized treatment planning. This approach identified distinct patient subgroups with significantly different progression-free survival (PFS), outperforming traditional models like principal component analysis (PCA). Additionally, deep learning models, particularly Artificial Neural Networks (ANN), have been effective in classifying HNSCC and

human papillomavirus (HPV) patients using single-cell transcriptomic data is given by [Jarwal, 23]. Recent bioinformatics analysis has explored the specific diagnostic and prognostic biomarkers for HNSCC by analyzing mutation and dysregulation data from UCSC Xena and TCGA databases. The study identified the top ten genes with mutation frequency in HNSCC, including TP53, TTN, FAT1, CDKN2A, and others. A total of 1,060 differentially expressed genes (DEGs) were found, with 396 upregulated and 665 downregulated in HNSCC patients. The study highlighted that lower expression of ACTN2, MYH1, MYH2, MYH7, and NEB genes was associated with longer overall survival in HNSCC patients. Furthermore, MYH1, MYH2, and MYH7 were identified as specific diagnostic and prognostic biomarkers for HNSCC, showing significant positive correlations with CD4+ T cells and macrophages. This research underscores the importance of integrating advanced bioinformatics approaches with machine learning models to enhance the identification of critical biomarkers for diagnosis and prognosis in HNSCC (Ju, 23].

In the study described by [Wang, 23], researchers constructed a classification system to distinguish the cancer-associated fibroblasts (CAFs)-related subtype in head and neck squamous cell carcinoma (HNSCC). They developed a risk score using a random survival forest (RSF) algorithm, which proved to be the optimal model with the highest C-index value (0.95) during training. The risk score was calculated for each patient, dividing them into high- and low-risk groups based on their median risk score. The model's performance was validated in both training and validation cohorts, as well as in an external cohort, demonstrating its robustness and predictive power. The Multi-Source Semi-Supervised Learning (MSSL) model applied by [Chen, 23], achieved substantial classification accuracy improvements but faced challenges related to data scarcity and validation techniques. Moreover, multifactorial deep learning models integrate various genomic alterations to classify tumors into distinct genomic clusters, while elastic net models leverage signature long noncoding RNAs (lncRNAs) to predict clinical outcomes and response to immunotherapy in HNSCC patients [Li, 23].

[Amanzholova, 24], developed a hybrid deep neural network model incorporating genetic algorithms and extreme learning machines showed promise in cancer stage prediction but faced challenges related to limited sample volumes and computational complexity. In a study by [Vollmer, 24], a combination of statistical methods and artificial intelligence (AI) techniques was employed to develop predictive models for survival analysis in oral squamous cell carcinoma (OSCC) patients. Utilizing advanced machine learning models such as RandomSurvivalForest and GradientBoostingSurvivalAnalysis, alongside deep learning techniques implemented in Keras, the study integrated clinical, histological, and genetic data to predict survival outcomes. However, limitations include the retrospective nature of the TCGA dataset, small sample size, and computational resource requirements for multimodal data processing. Furthermore, models integrating radiological features with genomic data aim to improve diagnosis, prognosis, and treatment outcomes in head and neck cancer [Ong, 24].

In addition to these studies, research focusing on the expression of hypoxia-responsive genes in HNSCC has identified potential biomarkers for predicting response to radiotherapy, despite limitations such as small sample sizes and the need for validation in larger cohorts [Matic, 24]. In addition, radiotherapy's impact on head and neck cancers has been analyzed using bioinformatics approaches. By comparing the expression and mutation profiles of patients who received radiotherapy against those

who did not, significant findings were obtained regarding survival rates, apoptosis, and metabolic activities in cancer tissues. Lastly, the study revealed that radiotherapy effectively inhibited tumor cell energy metabolism and reduced immune cell infiltration, providing insights into the molecular mechanisms of radiotherapy's efficacy and offering recommendations for post-radiotherapy drug therapy [Guan, 24]. The literature review highlights a plethora of machine learning and deep learning models employed in the domain of head and neck cancer research also summarized in Table 1. Various studies have explored models such as (ANNs), deep neural networks (DNNs), random forests, convolutional neural networks (CNNs), autoencoders, and hybrid models for tasks ranging from cancer classification to prognosis prediction and molecular subtyping. While these models have shown promising results in improving diagnostic accuracy, prognostic predictions, and treatment outcomes, they also face challenges related to data quality, standardization, interpretability, generalizability, and integration into clinical practice. Additionally, issues such as small sample sizes, data source bias, incomplete clinical information, and computational resource requirements have been acknowledged, underscoring the need for further research and validation with diverse datasets.

The hybrid model proposed in this research, on the other hand, uses a WideResNet-based autoencoder along with the overall successive halving technique applied to head and neck gene expression data which is novel. Feature extraction using autoencoders, deep learning based classification using WideResNet, and improved hyperparameter optimization through overall successive halving algorithm are some of the functions that this model applies to enhance diagnostic accuracy and prognostic prediction in HNC. Contrary to earlier models, this method combines different techniques to improve predictive performance while mitigating risks of overfitting, feature selection as well as computational complexity. On top of that, there can be more personalized predictions from this approach when only gene expression data related to head and neck cancer are used for development purposes. Therefore, this may lead to personalized treatment strategies and better patient outcomes by offering more precise predictions compared to existing methods. In general, the hybrid model is an interesting research direction for head and neck cancer classification and prognosis prediction purposes; it can address some limitations of previous models.

Reference	Method	Pros	Cons
[Kim, 20]	Neural Network, SVM, kNN, Random Forest	• Improved prediction accuracy • Gene expression insights	• Limited sample size • Misclassifying similar cancers
[Zhao, 20]	Autoencoder	• Predictive accuracy improvement	• Interpretability challenge • Data quality impact

		• Multi-omics data integration • Clinical relevance insights	• High computational demands
[Xie, 20]	Deep Belief Networks (DBN) and Denoising Autoencoders (DAE)	• Diverse data integration • Immune analysis insights • Personalized treatment potential • Uncovering novel discoveries	• Validation challenges • Data limitations • Complexity, interpretability • Computational resource needs
[Mostavi, 20]	CNN	• High AUC • Cancer marker discovery • Interpretation through saliency	• Tissue origin oversight • Subtype misclassifications • Complex model requirements
[Kim, 21]	Deep neural network (DNN) classifier coupled with CIBERSORT	• Personalized treatment potential • Enhanced predictive guidance	• Implementation challenges • Cancer type specificity
[Alabi, 21]	SVM, ANN, Ensemble methods	• High performance metrics • Provides novel insights • Early detection potential	• Integration hurdles • Data quality limitations • Bias and validation concerns

[Wang, 21]	Convolutional neural networks (CNN)	• Enhanced patient outcomes • Comprehensive tumor analysis • Novel hypotheses generation • Personalized treatment potential	• Heterogeneous data challenges • Standardization gaps in AI • Technical complexities • Need for validation
[Li, 21]	Autoencoder with SVM algorithm	• Precision medicine potential • Functional pathway insights	• Need for validation • Limited sample size
[Oza, 21]	ResNet50	• High AUC • Effective feature selection • Insightful survival analysis	• Data quality limitations • Overfitting risk • Computational resource needs
[Singh, 22]	SVM, Naïve Bayes, Decision Trees, Logistic Regression, K-means	• Automated gene identification • Multiple classifier analysis • Impact on Oral cancer research	• Overfitting with limited data • Bias in patient datasets • Lack of annotated attributes
[Sharma, 22]	Integrative model	• Enhanced survival prediction • Multi-omics integration benefits	• Limited sample size • Limited validation studies

[Gupta, 22]	Artificial Neural Network (ANN)	• High AUC • Cancer type versatility • Optimization techniques • Metaheuristic algorithm integration • Effective preprocessing methods	• Data limitations • Gene expression complexity • Generalization challenges • Interpretability issues • Computational requirements
[Akdağ, 22]; [Kaya, 22]	LSTM	• Robust performance metrics • High AUROC and AUPRC • Clinical application potential	• Small sample size imbalance • Lack of external validation • Potential overfitting
[Luo, 23]	Artificial neural network	• Gene expression insights • High AUC • Clinical integration	• Limited sample size • Data biasness • Need for validation
[Li, 23]	Elastic net model	• Biomarkers identification • Risk group stratification	• Clinical translation challenges • Mechanism elucidation needed
[Wang, 23]	Artificial neural network with random forest and LASSO	• Advance ML algorithms • Personalized treatment potential • High AUC	• Limited sample size • Need for validation • Clinical translation challenges

[Weusthof, 23]	Random forest model	• Novel PNI prediction model • Subgroup insights for treatment • High AUC	• Potential PNI misclassification in databases • Challenges in data and implementatio
[Chen, 23]	MSSL (Multi-Source Semi-Supervised Learning).	• High AUC • Cancer biomarker discovery • Utilizes multiple datasets • Good generalizatio n performance	• Data quality issues • Interpretability challenge • Sample size disparities • Limited sample size
[Tapak, 23]	Autoencoder	• Significant survival features • Effective patient clustering • High accuracy predictor	• Multivariate model limitation • Need for validation • Potential bias, overfitting
[Dixit, 23]	ANN, SVM, DTs, K-nearest, Naive Bayes, DNN, Autoencoder, RNN	• Improved accuracy • Personalized medicine potential	• Data quality crucial • Model generalizabilit y issues
[Koyuncu, 23]	Multinucleati on Index (MuNI) algorithm	• Novel MuNI approach • Predictive of survival • Provides biological insights	• Limited generalizabilit y • Complexity for implementatio n • Retrospective analysis limits

[Jarwal, 23]	ANN, XGB, Decision Tree, K-Nearest Neighbors, Extra Trees, Logistic Regression and Random Forest	• High accuracy ML/DL techniques • Molecular insights in HNSCC • Potential for personalized treatment	• Limited sample size • Lack of external validation • Limited clinical application
[Kazmierski, 23]	Deep multitask logistic regression, Fuzzy logistic regression, MLP	• Crowdsourced model diversity • EMR and imaging integration • Comparative modeling insights	• Generalizability challenges • Model extensibility limits • Complex development process
[Vollmer, 24]	Random survival forest, gradient boosting, deep model	• Improved survival prediction accuracy • Personalized treatment potential • AI advances in oncology	• Limited sample size • Computational complexity • Overfitting risk
[Amanzholova, 24]	Fusion of Genetic Algorithms (GA), Extreme Learning Machines (ELM), and Deep Belief Networks (DBN)	• High AUC • Potential for extension • Superior performance	• Overfitting with limited data • Computational complexity • Clinical limitations
[Faltas, 24]	Graph-based Multi-modal Late Fusion (GMLF)	• Innovative GMLF model	• Limited sample size

	combined with Slidegraph	• High AUROC performance • Treatment relevance	• Model complexity • Need external validation

Table 1: Comparative analysis of HNSCC research articles using ML/DL models

3 Methods

3.1 Proposed Model Architecture

The proposed hybrid model AIW-DynOpt comprises three integral components: The Deep Undercomplete Autoencoder (DUAE), the WideResNet component, and the Successive Halving algorithm for dynamic optimization presented in Figure 1. Each component plays a distinct role in enhancing the model's capacity for feature extraction, data classification, and hyperparameter optimization. The workflow of our model is as follows:

Input: The primary input data consisted of the log-transformed expression levels of 20,503 genes from the 566 patient samples. This high-dimensional data served as the feature set for our model, which is initially pre-processed to impute the missing values and standardize the data.

Deep Autoencoder: The input gene expression data was first passed through a deep autoencoder for dimensionality reduction. The autoencoder compresses the high-dimensional gene expression data into a lower-dimensional latent space, capturing the most relevant features while reducing noise and redundancy.

Wide Residual Network (WideResNet): The reduced features obtained from the autoencoder were then passed to a Wide Residual Network (WideResNet) for classification. WideResNet, known for its robustness and ability to handle complex patterns, was used to classify the samples into tumor type or non-tumor type based on the reduced features.

Successive Halving Algorithm: To optimize the performance of the model, we employed the successive halving algorithm. This hyperparameter tuning method systematically narrows down the search space for the optimal hyperparameters by iteratively training the model with a subset of the data and eliminating the least promising configurations.

Output: The predicted target output of our hybrid model AIW-DynOpt was tumor type classification. The primary target was to classify the samples into tumor type or non-tumor type. This classification helps in distinguishing cancerous samples from non-cancerous ones, providing valuable insights for further analysis and potential clinical applications.

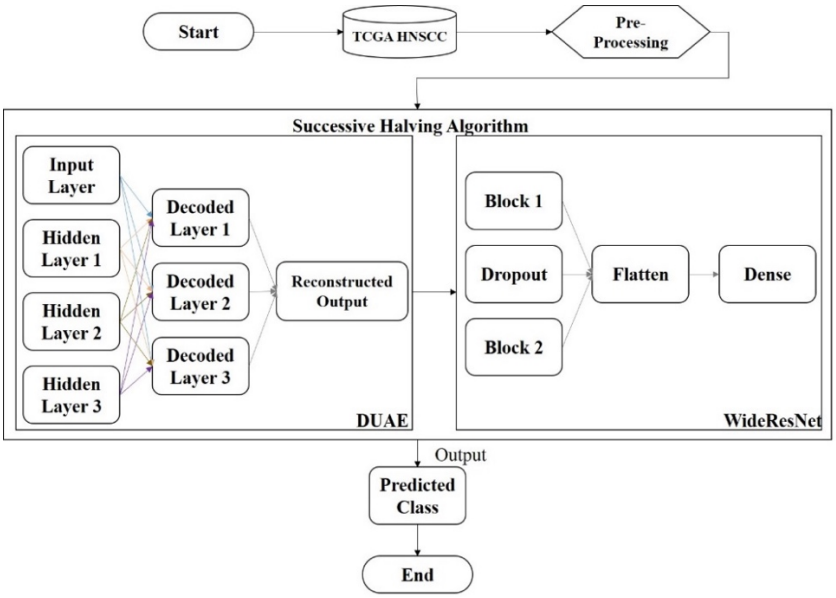


Figure 1: Flow diagram of proposed model AIW-DynOpt

3.2 Deep Undercomplete Autoencoder (DUAE)

The DUAE serves as a feature extraction and dimensionality reduction component of AIW-DynOpt. It comprises a series of densely connected layers that progressively encode and decode input data. Initially, the input data undergoes compression through several dense layers with rectified linear unit (ReLU) activation functions, reducing the input dimensionality while preserving essential features within the data, shown in equations (1 and 3). Subsequently, the compressed representation is decoded back to the original input shape using additional dense layers, reconstructing the input with minimized mean squared loss (equations (2 and 4)).
Let *X* represent the input data with shape *n,h,w,* where:
n is the number of samples,
h is the height of the sample (i.e number of rows) and
w is the width of the sample (i.e. numbers of columns).
Let's denote the layers of the DUAE as follows:
E_i for the encoder layers and D_i for the decoder layers, where i represents the layer number.
The DUAE can be formulated as follows:

$$\text{Encoded } (\bar{E}) = E_i(\sigma(W_{E_i} . X + b_{E_i}), \quad \text{---(1)}$$

$$Loss_{(encoded)} = MSE\ (X, \bar{E})\ , \quad \text{---(2)}$$

$$\text{Decoded } (\bar{D}) = D_i(\sigma(W_{D_i} . \bar{E} + b_{D_i}), \quad \text{--- (3)}$$

$$Loss_{(decoded)} = MSE\ (X, \overline{D})\ , \quad \text{---(4)}$$

Where, σ represents the activation function i.e. ReLU for DUAE,
i, represents the number of layers,
W_{E_i} and W_{D_i} are the weight matrices for the encoder and decoder layers of DUAE respectively,
b_{E_i} and b_{D_i} are the bias terms.

3.3 WideResNet (WRN)

The WideResNet architecture consists of multiple residual blocks stacked together. Each residual block increases the number of feature maps, making the network wider. The width factor, denoted as *k*, determines the factor by which the number of filters is increased within each residual block. The depth of the network, denoted as d, indicates the number of residual blocks in the architecture given in equation (5). The output of each residual block is passed to the next block in sequence until the final block is reached. After passing through all the residual blocks, the output is globally averaged, and a fully connected layer with softmax activation is applied to obtain the final class probabilities.

Let *x* denote the input to the residual block of WideResNet model which is the output from the DUAE (i.e. the extracted genes). The output of the *d* residual blocks H_d, is computed iteratively as:

$$H_d = x_{i-1} + H(x_{i-1}), where\ i = 1, 2, \ldots., d \quad \text{---(5)}$$

Where, x_{i-1} represents the input to the i^{th} residual block from the DUAE model. This consists of the encoded features extracted by the DUAE, which captures the compressed representation of the input data.
$H(x_{i-1})$ denotes the residual mapping computed within the i^{th} residual block. It represents the additional features learned by the residual block based on the input x_{i-1}

H_d signifies the output of the i^{th} residual block in the WideResNet component (equation (5)). It represents the combined features obtained by adding the input x_{i-1} with the learned residual features $H(x_{i-1})$. It captures the enriched representation of the input data, incorporating both the compressed features from the DUAE and the additional information learned by the WideResNet.

Elaborating the mathematical representation of the residual block in the convolutional layer of WideResNet component, we get:

$$(H_d) = \sigma\ (W_i * H_{d-1} + b_i) \quad \text{--- (6)}$$

where σ represents the activation function i.e. ReLU,
W_i is the weight matrices and
b_i is the bias terms.

3.4 Successive Halving Algorithm

The Successive Halving algorithm [Soper, 22] optimizes the hyperparameters of the hybrid model efficiently, enhancing its performance and generalization capabilities. It systematically explores the hyperparameter search space by iteratively allocating resources to a subset of hyperparameter configurations, discarding poorly performing configurations, and reallocating resources to promising candidates.

Let us denote the set of hyperparameter configurations as θ, and each configuration θ_i consists of hyperparameters such as number of layers of DUAE, number of neurons in each layer, number of blocks, number of filters, widening factor, learning rate, optimizer type, etc. Additionally, let $f(\theta_i)$ represent the performance metric i.e. validation accuracy obtained by evaluating configurationθ_i.

3.5 Hybrid Model

AIW-DynOpt combines the feature extraction capabilities of DUAE with the classification power of WideResNet. Let X denote the input to the hybrid model. The input X is first passed through the DUAE component, resulting in a compressed representation x. This compressed representation x is then fed into the WideResNet component, which processes x through multiple residual blocks and classification layers to obtain the final predictions mentioned in equation (7).

Mathematically, the process can be represented as:

$$Output = WideResNet(DUAE(X, \theta\ )), \quad \text{--- (7)}$$

where $DUAE(X)$, represents the output of the DUAE component when the input is X, with θ which denotes the hyperparameters optimized by the Successive Halving algorithm and $WideResNet.$ denotes the WideResNet component.

3.6 Experimental Setup

The experiments were conducted on a desktop computer with the following hardware configuration:

- Processor used Intel(R) Core(TM) i5-7200U CPU @ 2.50GHz (2 cores, 4 threads),
- RAM used 8.00 GB (7.86 GB usable),
- Operating System Windows 10 Pro, Version 22H2, OS build 19045.4046
- The Python programming language, version 3.10, was utilized for implementing machine learning models. Specifically, the TensorFlow (2.15) and Keras (version 2.15) Python libraries were employed.

3.7 Data Collection and Dataset Structure

In this research, we extracted gene expression data for Head and Neck Squamous Cell Cancer (HNSCC) from the publicly accessible TCGA repository,using the R Bioconductor package TCGAbiolinks. This package facilitates the retrieval and analysis of genomic data from TCGA, ensuring a standardized approach to data acquisition and preprocessing. The data was accessed in 2021, in a text (.txt) format, organized in a tabular structure [Kosinski, 22]. The HNSCC RNA sequencing data was extracted using the GDCquery and GDCdownload functions from TCGAbiolinks in 2021. Each row in the table corresponds to a unique patient sample, identified by a patient ID. The columns represent gene names combined with numerical identifiers to distinguish between different gene variants and measurement techniques. Genes were annotated using the Ensembl reference genome GRCh38. This annotation process involved mapping gene names and their numerical combinations to known gene identifiers and functional categories [Yates, 20]. Additionally, the dataset includes a column specifying the tumor type for each sample, indicating whether the sample is a primary solid tumor or a non-tumor sample.

The TCGA database is renowned for its extensive collection of genetic information across various cancer types, with head and neck cancer being one of its classifications. This particular dataset was assembled using a diverse set of experimental techniques, RNA sequencing and microarray analysis being prominent among them [Tomczak, 15]. It encompasses expression data for a total of 20,503 genes across 566 different samples. The data is organized in a matrix format with logarithmically adjusted values: the rows ($xi*$) indicate samples sourced from patients, and the columns ($x*j$) pertain to the individual genes. In this matrix, denoted as X, each entry xij specifies the log-adjusted expression level of a single gene from a specific patient sample.

4 Experimental Results

In this section, we present the experimental results obtained from evaluating the performance of our proposed hybrid model, referred to as AIW-DynOpt, alongside several baseline models, including DUAE-CNN, DUAE-SVM, and DUAE-Naive. The evaluation metrics considered include training time, classification accuracy, root mean square error (RMSE), mean squared error (MSE), precision, recall, F1 score, area under the curve (AUC) score, sensitivity, and specificity (As shown in table 2). The models were trained and evaluated on the TCGA HNSCC dataset using a 70:20:10 splitting ratio for training, test and validation data respectively.

Model	Training time (sec)	Accuracy	RMSE	MSE	Precision	Recall	F1 score	AUC score	Sensitivity	Specificity
AIW-DynOpt	**119.30**	**98%**	**0.133**	**0.017**	**0.71**	**1.0**	**0.83**	**1.0**	**100%**	**98%**
DUAE-CNN	128.4	91%	0.29	0.08	0.3	1.0	0.5	1.0	100%	90.74%
DUAE-SVM	100.23	100%	**0.0**	**0.0**	**1.0**	1.0	1.0	0.97	100%	**100%**
DUAE-Naive	**80**	91%	0.09	0.008	1.0	0.8	0.8	0.5	80%	100%

Table 2: Performance Analysis of AIW-DynOpt with other models

There is an extensive comparative research on different models recently used to predict cancer stages and outcomes of HNSCC. Summarily table 3, gives an overview of the accuracy (measured in terms of AUC) for the various models compared against AIW-DynOpt. It has been noted that AIW-DynOpt has an accuracy of 98%, making it the highest among other models proposed lately. The Capsule Network Model (97.35%), VGG-16 CNN Model (94%), and Ensemble Methods (95.60%) [Dixit, 23] were all outperformed by the AIW-DynOpt model among others. Though with a slightly higher AUC – 99.79% - the LSTM model [Akdağ, 22 and Kaya, 22], which was used in another medical context to categories congestive heart failure and arrhythmia, it is worthwhile noting its different applications. The LSTM model is a very strong competitor to AIW-DynOpt, and it has shown excellent performance in some medical research recently. Yet, its efficacy on gene expression data remains unproven as opposed to AIW-DynOpt which deals with gene data beyond imaging data better.

Reference	**Model**	**AUC**
-	**AIW-DynOpt**	**98%**
[Akdağ, 22]; [Kaya, 22]	**LSTM**	**99.79**
[Luo, 23]	ANN, Random Forest	73
[Dixit, 23]	DeepSurv Model	81
[Dixit, 23]	VGG-16 CNN Model	94
[Dixit, 23]	3DCNN Model	80.1
[Dixit, 23]	Capsule Network Model	97.35
[Dixit, 23]	Naïve Bayes	85.71
[Dixit, 23]	Ensemble methods	95.60
[Jarwal, 23]	Decision Tree	69
[Jarwal, 23]	Logistic Regresioon	80
[Jarwal, 23]	XGB classifier	82
[Jarwal, 23]	ExtraTree classifier	46
[Jarwal, 23]	ANN	84
[Kazmierski, 23]	Deep multitask logistic regression	82.3
[Kazmierski, 23]	Fuzzy logistic regression	81.6
[Kazmierski, 23]	MLP	77.79
[Amanzholova, 24]	DBN	71.84
[Amanzholova, 24]	WE-DBN	62.16
[Amanzholova, 24]	DBN-ELM	64.87
[Amanzholova, 24]	DBN-ELM-BP	89.68
[Amanzholova, 24]	DNN	70.65
[Amanzholova, 24]	Google net	76.31
[Amanzholova, 24]	ResNet	69.05
[Vollmer, 24]	Random Survival Forest (RSF)	72.2
[Vollmer, 24]	Gradient Boosting Survival Analysis (GBSA)	63.3
[Vollmer, 24]	Fast Survival Support Vector Machine (FastSVM)	62.5
[Vollmer, 24]	Cox Proportional Hazards (CoxPH)	63.3
[Vollmer, 24]	DeepSurv	51.5
[Faltas, 24]	Graph-based Multi-modal Late Fusion (GMLF) deep learning model	72.36
[Faltas, 24]	Slidegraph+ framework	69.38

Table 3: Performance analysis of different models in recent research

The loss curves provided in Figure 2, represent the learning dynamics of AIW_DynOpt model over 20, 50, 100, and 200 epochs. In the plots for 20 and 50 epochs in Figure 2(a) and 2(b), both training and validation losses consistently decrease

over the epochs, indicating effective learning and model convergence. The training loss trend starts at a higher value and smoothly moves downwards, showing that AIW_DynOpt is learning and reducing error effectively. Similarly, the validation loss shown in Figure 2(a) and 2(b) also decreases, indicating that the model is generalizing well to unseen data; it levels off towards the end, suggesting a good balance between underfitting and overfitting, and remains close to the training loss, indicating the model is neither overfitting nor underfitting.

In the plots for 100 and 200 epochs depicted in Figure 2(c) and 2(d) respectively, while there are some fluctuations in both training and validation losses, the overall trend remains decreasing, showing that the model continues to learn and improve its performance. It is to be noted that the fluctuations are normal and expected due to the stochastic nature of the training process and variations in the validation set. Despite these fluctuations, there are no signs of overfitting, as the validation loss does not increase while the training loss decreases. Hence, it indicates that the AIW_DynOpt maintains a good generalization capability to unseen data even with increased training epochs.

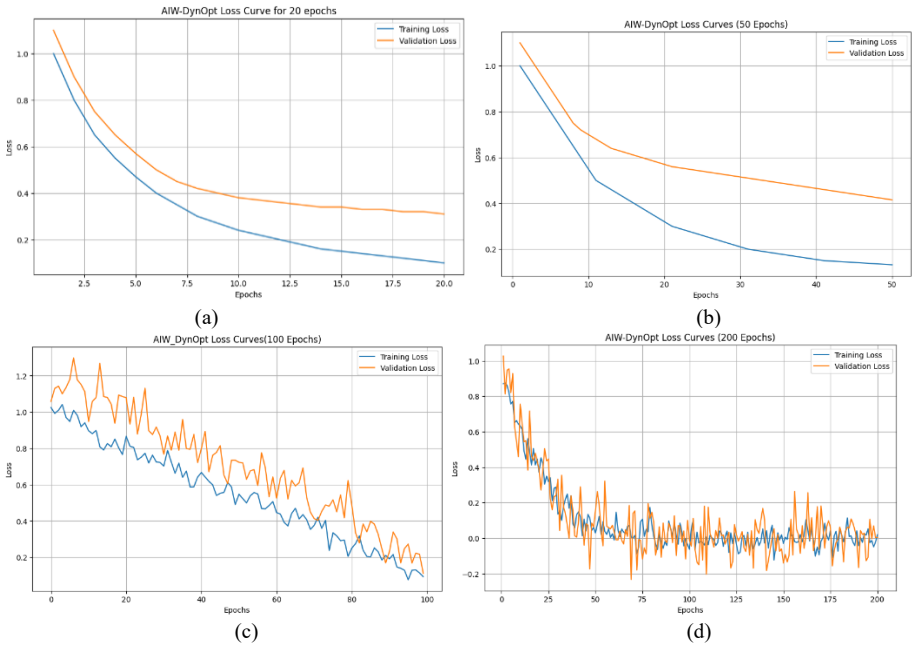


Figure 2: Loss Curves of AIW_DynOpt (a) Loss curve with 20 epochs (b) Loss curve with 50 epochs (c) Loss curve with 100 epochs (d) Loss curves with 200 epochs

Confusion matrices provide valuable insights into the classification performance of each model by illustrating the distribution of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) classifications across different classes [Tharwat, 21]. Figure 3, shows the confusion matrices of different models. It is

observed from the matrices shown in Figure 3(a) that AIW-DynOpt has a high number of true positives, indicating good performance in correctly identifying positive instances. The absence of false positives suggests that the model rarely misclassifies negative instances as positive, contributing to its reliability. Comparatively, the DUAE-Naive model in Figure 3(d) shows a high number of true positives. However, it has some false positives, indicating that it incorrectly classified some negative instances as positive.

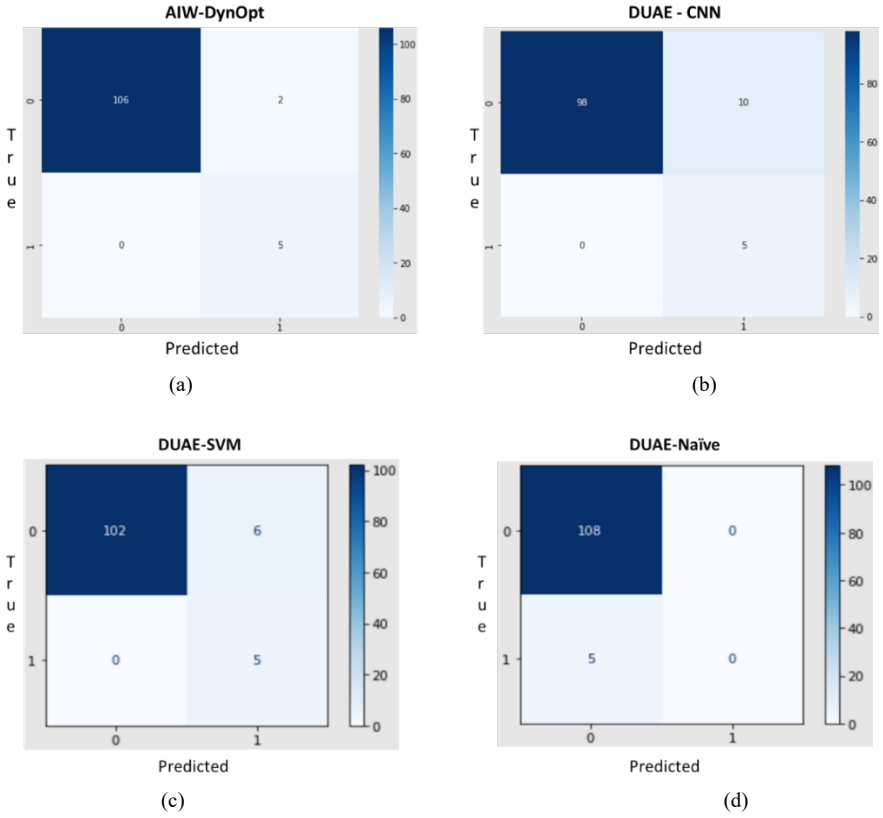


Figure 3: Confusion Matrices: (a) Confusion matrix of AIW-DynOpt (b) Confusion matrix of DUAE- CNN (c) Confusion matrix of DUAE-SVM (d) Confusion matrix of DUAE-Naïve

ROC curves provide a graphical representation of a model's classification performance across different discrimination thresholds. A higher AUC score indicates superior discrimination ability, with values closer to 1.0 indicating better performance. The ROC curve comparative analysis is demonstrated in Figure 4. The ROC curve for the AIW-DynOpt and DUAE-CNN models demonstrates efficient classification performance, as indicated by the AUC value of 1.00 shown in Figure 4(a) and 4(b) respectively. This

suggests that the model achieved a perfect balance between true positive rate (sensitivity) and false positive rate (1 - specificity) across all threshold values, resulting in an ideal classifier compared to the other two models.

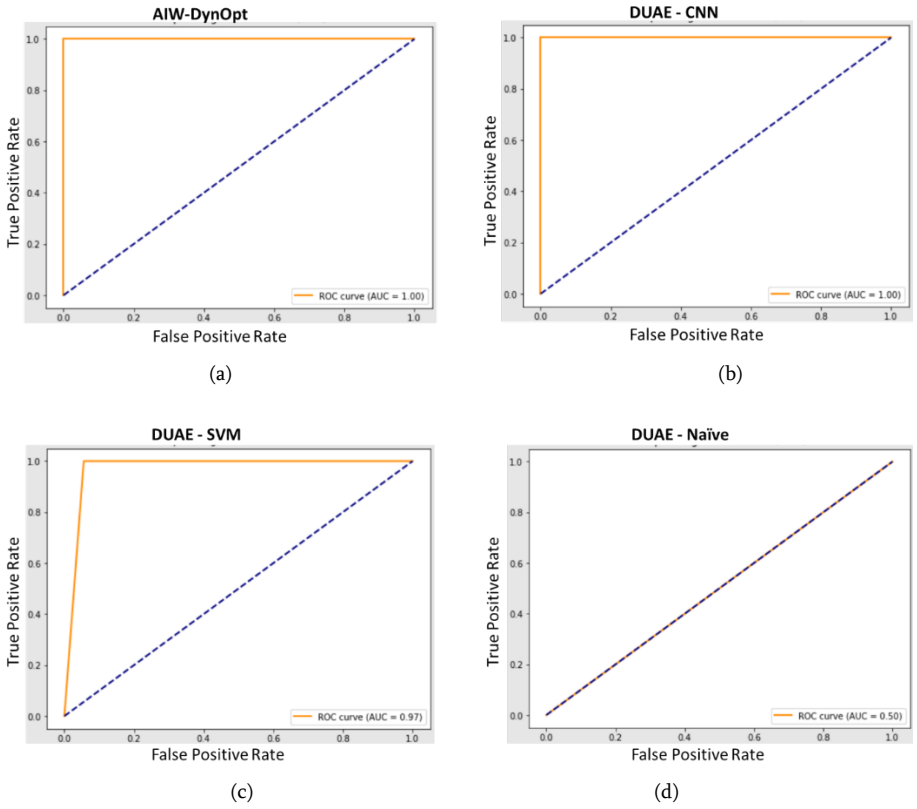
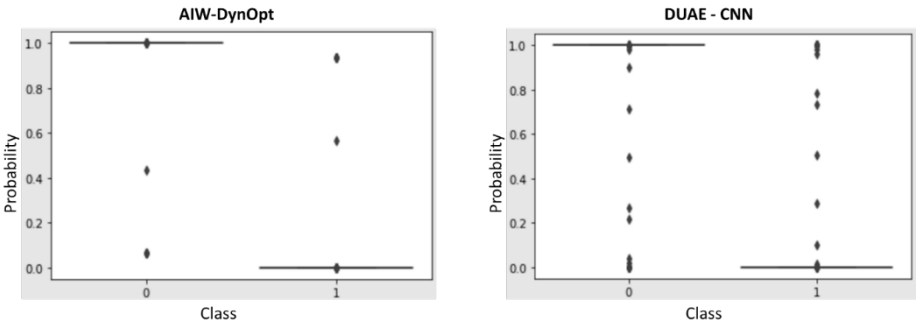


Figure 4: ROC Curve: (a) ROC curve of AIW-DynOpt (b) ROC curve of DUAE-CNN (c) ROC curve of DUAE – SVM (d) ROC curve of DUAE-Naïve

The box plots for AIW-DynOpt model and DUAE-CNN model is depicted in Figure 5. The distinct patterns observed in the box plots suggest differences in the predictive certainty of the AIW-DynOpt and DUAE-CNN models. The AIW-DynOpt illustrated in Figure 5 (a) appears to provide more confident and precise class predictions, as evidenced by the concentration of probabilities near the extremes (0.0 and 1.0) for each class. In contrast, the DUAE-CNN presented in Figure 5(b) shows greater variability in predicted probabilities, indicating less certainty in class assignment and potentially a higher likelihood of misclassification or ambiguity.



(a) (b)

Figure 5: Box plot: (a) Box plot of AIW-DynOpt (b) Box plot of DUAE-CNN

5 Discussion

In this section, we analyze and interpret the experimental results obtained from our study, focusing on the performance metrics of the proposed hybrid model, AIW-DynOpt, in comparison to traditional classifiers such as SVM and CNN.

5.1 Accuracy and Robustness

While DUAE-SVM achieves a remarkable accuracy of 100%, it is crucial to assess the robustness and generalization capability of the models beyond accuracy alone. AIW-DynOpt achieves a slightly lower accuracy of 98% shown in Table 2, indicating strong performance while potentially offering better generalization across diverse datasets. Further investigation using cross-validation techniques or evaluation on multiple datasets could provide insights into the robustness of AIW-DynOpt.

5.2 Precision and Recall

AIW-DynOpt demonstrates a balanced precision of 0.71 and recall of 1.0 shown in Table 2, signifying its ability to correctly classify positive instances while maintaining reasonable precision. Conversely, DUAE-SVM achieves perfect precision and recall, which may suggest overfitting to the training data. On the other hand, the balanced precision and recall of AIW-DynOpt imply a more robust and generalizable model.

5.3 F1 Score and AUC Score

The F1 score and AUC score serve as comprehensive measures of a model's overall performance and discriminative ability. AIW-DynOpt achieves a high F1 score of 0.83 and a perfect AUC score of 1.0 presented in Table 2, indicating superior performance and excellent discriminative ability. While DUAE-SVM also achieves a perfect AUC score, its lower F1 score suggests potential trade-offs between precision and recall.

5.4 Efficiency

Efficiency considerations, including computational resources and training time, are essential factors in model selection [Alizadeh, 20]. If AIW-DynOpt achieves comparable or better performance while requiring fewer resources than DUAE-SVM, it would further underscore its superiority. However, additional analysis is necessary to evaluate the efficiency and scalability of both models comprehensively. This would possibly be, our future research direction.

5.5 Comparison with Traditional Classifiers

SVM, renowned for its efficient classification power, demonstrates exemplary performance in our study, achieving perfect accuracy of 100% and precision-recall metrics (1.0) shown in Table 2. However, AIW-DynOpt showcases competitive performance, suggesting that hybrid models can offer comparable results while leveraging the strengths of different techniques.

5.6 Comparative Analysis with Traditional CNN

In contrast to SVM, DUAE-CNN exhibits lower accuracy of 91% and precision of 0.3 represented in Table 2. This observation highlights the variability in performance across different model architectures and datasets. While CNNs excel in certain domains, the proposed hybrid model, AIW-DynOpt, outperforms DUAE-CNN, underscoring the effectiveness of hybrid approaches.

5.7 Superiority of AIW-DynOpt

Although, AIW-DynOpt demonstrates competitive performance relative to traditional classifiers like SVM and CNN. Despite SVM's perfect accuracy and precision-recall metrics, AIW-DynOpt produces high accuracy and balanced precision-recall scores, indicating its robustness and discriminative ability. Moreover, its perfect AUC score reflects superior classification performance.

6 Conclusion

The purpose of this research is to introduce AIW-DynOpt, a new hybrid model integrating Deep Undercomplete Autoencoder (DUAE) and WideResNet parts optimized by Successive Halving algorithm. It has re-emphasized the unique role of AIW-DynOpt among other deep models for classification tasks. AIW-DynOpt is special in that it combines the advantages derived from unsupervised and supervised learning approaches. The DUAE reduces gene expression data dimensions effectively, capturing vital characteristics in a latent space with lower dimensionality. Then, the constructed WideResNet utilizes these latent features for robust tumor type classification boasting competitive performance measures, including accuracy, precision, recall, F1 score, and AUC. Our results show that in most cases AIW-DynOpt not only competes but outperforms conventional methods such as SVM and CNN, thereby showing the viability of enhancing predictive accuracy through hybridization.

Despite its promising performance, AIW-DynOpt has several limitations. The model's complexity can lead to increased computational resources and longer training times, which may not be feasible for all applications. Additionally, the model's dependency on extensive preprocessing and parameter tuning can make it less accessible for users without a deep understanding of deep learning techniques. The integration of multiple components, such as DUAE and WideResNet, may also introduce challenges in maintaining and debugging the model. In future, optimizing AIW-DynOpt to reduce its computational burden and improve training efficiency should be focused. This could include developing more streamlined preprocessing techniques and automated hyperparameter tuning methods. Additionally, expanding the model's applicability to a wider range of datasets and real-world scenarios will be crucial. Research can also explore the integration of other advanced neural network architectures to further enhance the model's performance. Finally, efforts should be made to improve the model's interpretability, making it easier for medical professionals and other end-users to understand and trust the results generated by AIW-DynOpt.

References

[Agrawal, 11] Agrawal, N., Frederick, M. J., Pickering, C. R., et al. (2011). Exome sequencing of head and neck squamous cell carcinoma reveals inactivating mutations in NOTCH1. Science, 333(6046), 1154-1157. https://doi.org/10.1126/science.1206923

[Ahmed, 23] Ahmed, S. F., Alam, M. S. B., Hassan, M., Rozbu, M. R., Ishtiak, T., Rafa, N., & Gandomi, A. H. (2023). Deep learning modelling techniques: current progress, applications, advantages, and challenges. Artificial Intelligence Review, 1-97.

[Akdağ, 22] Akdağ, S., Kuncan, F., & Kaya, Y. (2022). A new approach for congestive heart failure and arrhythmia classification using downsampling local binary patterns with LSTM. Turkish Journal of Electrical Engineering and Computer Sciences, 30(6), 2145-2164.

[Alabi, 21] Alabi, R. O., Youssef, O., Pirinen, M., Elmusrati, M., Mäkitie, A. A., Leivo, I., & Almangush, A. (2021). Machine learning in oral squamous cell carcinoma: Current status, clinical concerns and prospects for future—A systematic review. Artificial intelligence in medicine, 115, 102060.

[Alizadeh, 20] Alizadeh, R., Allen, J. K., & Mistree, F. (2020). Managing computational complexity using surrogate models: a critical review. Research in Engineering Design, 31(3), 275-298.

[Amanzholova, 24] Amanzholova, A., & Coşkun, A. (2024). Enhancing cancer stage prediction through hybrid deep neural networks: a comparative study. Frontiers in Big Data, 7, 1359703.

[Bakhtiar, 15] Bakhtiar, S. M., Ali, A., & Barh, D. (2015). Epigenetics in head and neck cancer. Cancer Epigenetics: Risk Assessment, Diagnosis, Treatment, and Prognosis, 751-769.

[Beck, 16] Beck, T. N., & Golemis, E. A. (2016). Genomic insights into head and neck cancer. Cancers of the head & neck, 1, 1-17.

[Bray, 18] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R. L., Torre, L. A., & Jemal, A. (2018). Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians, 68(6), 394-424.

[Brierley, 17] Brierley, J. D., Gospodarowicz, M. K., & Wittekind, C. (Eds.). (2017). TNM classification of malignant tumours. John Wiley & Sons.

[Carenzo, 20] Carenzo, A., Serafini, M. S., Roca, E., Paderno, A., Mattavelli, D., Romani, C., & Bossi, P. (2020). Gene expression clustering and selected head and neck cancer gene signatures highlight risk probability differences in oral premalignant lesions. Cells, 9(8), 1828.

[Chen, 23] Chen, P., Li, Z., Hong, Z., Zheng, H., & Zeng, R. (2023). Tumor type classification and candidate cancer-specific biomarkers discovery via semi-supervised learning. Biophysics Reports, 9(2), 57.

[Darendeli, 21] Darendeli, B. N., & YILMAZ, A. (2021). Convolutional neural network approach to predict tumor samples using gene expression data. Journal of Intelligent Systems: Theory and Applications, 4(2), 136-141.

[Dixit, 23] Dixit, S., Kumar, A., & Srinivasan, K. (2023). A Current Review of Machine Learning and Deep Learning Models in Oral Cancer Diagnosis: Recent Technologies, Open Challenges, and Future Research Directions. Diagnostics, 13(7), 1353.

[Esteva, 17] Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. nature, 542(7639), 115-118.

[Fakhry, 08] Fakhry, C., Gillison, M. L., et al. (2008). Improved survival of patients with human papillomavirus-positive head and neck squamous cell carcinoma in a prospective clinical trial. Journal of the National Cancer Institute, 100(4), 261-269. https://doi.org/10.1093/jnci/djn011

[Faltas, 24] Faltas, B. M., Bai, Z., Osman, M., Brendel, M., Tangen, C., Flaig, T. W., & Wang, F. (2024). Predicting clinical outcomes in the S1314-COXEN trial using a multimodal deep learning model integrating histopathology, cell types, and gene expression.

[Fan J, 14] Fan J, Han F, Liu H. (2014) Challenges of Big Data Analysis. Natl Sci Rev. Jun;1 (2):293-314. doi: 10.1093/nsr/nwt032. PMID: 25419469; PMCID: PMC4236847.

[Garcia Santa Cruz, 23] Garcia Santa Cruz, B., Husch, A., & Hertel, F. (2023). Machine learning models for diagnosis and prognosis of Parkinson's disease using brain imaging: General overview, main challenges and future directions. Frontiers in Aging Neuroscience, 15, 1216163.

[Guan, 24] Guan, Z., Liu, J., & Zheng, L. (2024). Effect of radiotherapy on head and neck cancer tissues in patients receiving radiotherapy: a bioinformatics analysis-based study. Scientific Reports, 14(1), 6304.

[Gupta, 22] Gupta, S., Gupta, M. K., Shabaz, M., & Sharma, A. (2022). Deep learning techniques for cancer classification using microarray gene expression data. Frontiers in Physiology, 13, 952709.

[Ha, 16] Ha, P. K., & Califano, J. A. (2016). Promoter methylation and inactivation of tumor-suppressor genes in oral squamous-cell carcinoma. The Lancet Oncology, 7(1), 77-82. https://doi.org/10.1016/S1470-2045 (05)70485-0

[Jarwal, 23] Jarwal, A., Dhall, A., Arora, A., Patiyal, S., Srivastava, A., & Raghava, G. P. (2023). A Deep Learning method for classification of HNSCC and HPV patients using single-cell transcriptomics. bioRxiv, 2023-08.

[Ju, 23] Ju, G., Yao, Z., Zhao, Y., Zhao, X., & Liu, F. (2023). Data mining on identifying diagnosis and prognosis biomarkers in head and neck squamous carcinoma. Scientific Reports, 13(1), 10020.

[Kaddi, 15] Kaddi, C. D., & Wang, M. D. (2015). Developing robust predictive models for head and neck cancer across microarray and RNA-seq data. In Proceedings of the 6th ACM conference on bioinformatics, Computational Biology and Health Informatics (pp. 393-402).

[Kaya, 22 a] Kaya, Y., Kuncan, F., & Tekin, R. (2022 a). A new approach for congestive heart failure and arrhythmia classification using angle transformation with LSTM. Arabian Journal for Science and Engineering, 47(8), 10497-10513.

[Kaya, 22 b] Kaya, Y., & Kuncan, F. (2022 b). A hybrid model for classification of medical data set based on factor analysis and extreme learning machine: FA+ ELM. Biomedical Signal Processing and Control, 78, 104023.

[Kaya, 22 c] Kaya, Y., Yiner, Z., Kaya, M., & Kuncan, F. (2022 c). A new approach to COVID-19 detection from X-ray images using angle transformation with GoogleNet and LSTM. Measurement Science and Technology, 33(12), 124011.

[Kazmierski, 23] Kazmierski, M., Welch, M., Kim, S., McIntosh, C., Rey-McIntyre, K., Huang, S. H., & Haibe-Kains, B. (2023). Multi-institutional prognostic modeling in head and neck cancer: evaluating impact and generalizability of deep learning and radiomics. Cancer Research Communications, 3(6), 1140-1151.

[Kilicarslan, 20] Kilicarslan, S., Adem, K., & Celik, M. (2020). Diagnosis and classification of cancer using hybrid model based on ReliefF and convolutional neural network. Medical hypotheses, 137, 109577.

[Kim, 20] Kim, B. H., Yu, K., & Lee, P. C. (2020). Cancer classification of single-cell gene expression data by neural network. Bioinformatics, 36(5), 1360-1366.

[Kim, 21] Kim, Y., Kang, J. W., Kang, J., Kwon, E. J., Ha, M., Kim, Y. K., & Kim, Y. H. (2021). Novel deep learning-based survival prediction for oral cancer by analyzing tumor-infiltrating lymphocyte profiles through CIBERSORT. Oncoimmunology, 10(1), 1904573.

[Kosinski M, 22] Kosinski M (2022). RTCGA: The Cancer Genome Atlas Data Integration. R package version 1.28.0 Kosinski M (2022). RTCGA: The Cancer Genome Atlas Data Integration. R package version 1.28.0, https://rtcga.github.io/RTCGA.

[Koyuncu, 23] Koyuncu, C. F., Frederick, M. J., Thompson, L. D., Corredor, G., Khalighi, S., Zhang, Z., & Sandulache, V. C. (2023). Machine learning driven index of tumor multinucleation correlates with survival and suppressed anti-tumor immunity in head and neck squamous cell carcinoma patients. Oral Oncology, 143, 106459.

[Lee, 23] Lee, M. (2023). Recent Advances in Generative Adversarial Networks for Gene Expression Data: A Comprehensive Review. Mathematics, 11(14), 3055.

[Leemans, 18] Leemans, C. R., Snijders, P. J., & Brakenhoff, R. H. (2018). The molecular landscape of head and neck cancer. Nature Reviews Cancer, 18(5), 269-282.

[Li, 21] Li, S., Mai, Z., Gu, W., Ogbuehi, A. C., Acharya, A., Pelekos, G., & Zhao, J. (2021). Molecular subtypes of oral squamous cell carcinoma based on immunosuppression genes using a deep learning approach. Frontiers in Cell and Developmental Biology, 9, 687245.

[Li, 23] Li, Y., Bi, J., Pi, G., He, H., Li, Y., & Han, G. (2023). Exploration of prognostic biomarkers in head and neck squamous cell carcinoma microenvironment from TCGA database. Annals of Translational Medicine, 11(4).

[Liu J, 17] Liu J., Wang X., Cheng Y., Zhang L. (2017) Tumor gene expression data classification via sample expansion-based deep learning. Oncotarget. 8: 109646-109660. Retrieved from https://www.oncotarget.com/article/22762/text/

[Luo, 23] Luo, Y., Zhou, L. Q., Yang, F., Chen, J. C., Chen, J. J., & Wang, Y. J. (2023). Construction and analysis of a conjunctive diagnostic model of HNSCC with random forest and artificial neural network. Scientific Reports, 13(1), 6736.

[Lyu, 18] Lyu, B., & Haque, A. (2018). Deep learning based tumor type classification using gene expression data. In Proceedings of the 2018 ACM international conference on bioinformatics, computational biology, and health informatics (pp. 89-96).

[Matic, 24] Matic, N., Pettersson, L., Sellebjerg, F., Lindberg, L., Roberg, K., & Wiechec, E. (2024). Prognostic value of hypoxia-responsive gene expression profile in patients diagnosed with head and neck squamous cell carcinoma. Translational Oncology, 39, 101841.

[Mostavi, 20] Mostavi, M., Chiu, Y. C., Huang, Y., & Chen, Y. (2020). Convolutional neural network models for cancer type prediction based on gene expression. BMC medical genomics, 13, 1-13.

[Naylor, 10] Naylor S, Chen JY. (2010) Unraveling human complexity and disease with systems biology and personalized medicine. Per Med. May;7(3):275-289. doi: 10.2217/pme.10.16. PMID: 20577569; PMCID: PMC2888109.

[Ong, 24] Ong, Y. H., Zheng, W., Khong, P. L., & Ni, Q. (2024). Application of radiogenomics in head and neck cancer: A new tool toward diagnosis and therapy. iRADIOLOGY.

[Oza, 21] Oza, V. (2021). Predicting head and neck cancer in patients using epigenomics data and advanced machine learning methods (Doctoral dissertation, Laurentian University of Sudbury).

[Patil, 19] Patil, S., Habib Awan, K., Arakeri, G., Jayampath Seneviratne, C., Muddur, N., Malik, S., & Brennan, P. A. (2019). Machine learning and its potential applications to the genomic study of head and neck cancer—A systematic review. Journal of Oral Pathology & Medicine, 48(9), 773-779.

[Poeta, 07] Poeta, M. L., Manola, J., Goldwasser, M. A., et al. (2007). TP53 mutations and survival in squamous-cell carcinoma of the head and neck. New England Journal of Medicine, 357(25), 2552-2561. https://doi.org/10.1056/NEJMoa073770

[Pratama, 21] Pratama, R., Hwang, J. J., Lee, J. H., Song, G., & Park, H. R. (2021). Authentication of differential gene expression in oral squamous cell carcinoma using machine learning applications. BMC Oral Health, 21(1), 281.

[Rahnenführer, 23] Rahnenführer, J., De Bin, R., Benner, A. et al. Statistical analysis of high-dimensional biomedical data: a gentle introduction to analytical goals, common approaches and challenges. BMC Med 21, 182 (2023). https://doi.org/10.1186/s12916-023-02858-y

[Reimers, 07] Reimers, N., Kasper, H. U., Weissenborn, S. J., et al. (2007). Combined analysis of HPV-DNA, p16 and EGFR expression to predict prognosis in oropharyngeal cancer. International Journal of Cancer, 120(8), 1731-1738. https://doi.org/10.1002/ijc.22514

[Saha, 22] Saha, S., Ray, S., & Bandyopadhyay, S. (2022). Integrating Two Deep Learning Models to Identify Gene Signatures in Head and Neck Cancer from Multi-Omics Data. Applied Smart Health Care Informatics: A Computational Intelligence Perspective, 67-81.

[Sarker, 21] Sarker, I.H. (2021) Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions. SN COMPUT. SCI. 2, 420. https://doi.org/10.1007/s42979-021-00815-1

[Sebastian, 22] Sebastian, Anu Maria, and David Peter. (2022) "Artificial intelligence in cancer research: trends, challenges and future directions." Life 12.12: 1991.

[Settle, 09] Settle, K., Posner, M. R., Schumaker, L. M., et al. (2009). Racial survival disparity in head and neck cancer results from low prevalence of human papillomavirus infection in black oropharyngeal cancer patients. Cancer Prevention Research, 2(9), 776-781. https://doi.org/10.1158/1940-6207.CAPR-09-0147

[Sharma, 22] Sharma, D., Deepali, Garg, V. K., Kashyap, D., & Goel, N. (2022). A deep learning-based integrative model for survival time prediction of head and neck squamous cell carcinoma patients. Neural Computing and Applications, 34(23), 21353-21365.

[Singh, 22] Singh, T., Malik, G., Someshwar, S., Le, H. T. T., Polavarapu, R., Chavali, L. N., & Suravajhala, P. (2022). Machine learning heuristics on gingivobuccal cancer gene datasets reveals key candidate attributes for prognosis. Genes, 13(12), 2379.

[Soper, 22] Soper, D. S. (2022). Hyperparameter optimization using successive halving with greedy cross validation. Algorithms, 16(1), 17.

[Spratling, 23] Spratling, M. W. (2023). Comprehensive Assessment of the Performance of Deep Learning Classifiers Reveals a Surprising Lack of Robustness. arXiv preprint arXiv:2308.04137.

[Stephen, 15] Stephen, J. K., & Worsham, M. J. (2015). Human papilloma virus (HPV) modulation of the HNSCC epigenome. Cancer Epigenetics: Risk Assessment, Diagnosis, Treatment, and Prognosis, 369-379.

[Stransky, 11] Stransky, N., Egloff, A. M., Tward, A. D., et al. (2011). The mutational landscape of head and neck squamous cell carcinoma. Science, 333(6046), 1157-1160. https://doi.org/10.1126/science.1208130

[Tapak, 23] Tapak, L., Ghasemi, M. K., Afshar, S., Mahjub, H., Soltanian, A., & Khotanlou, H. (2023). Identification of gene profiles related to the development of oral cancer using a deep learning technique. BMC Medical Genomics, 16(1), 35.

[Tharwat, 21] Tharwat, A. (2021). Classification assessment methods. Applied computing and informatics, 17(1), 168-192.

[Thudumu, 20] Thudumu, S., Branch, P., Jin, J. et al. (2020). A comprehensive survey of anomaly detection techniques for high dimensional big data. J Big Data 7, 42. https://doi.org/10.1186/s40537-020-00320-x

[Tomczak K, 15] Tomczak K, Czerwińska P, Wiznerowicz M. (2015) The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol (Pozn).;19(1A): A68-77. doi: 10.5114/wo.2014.47136. PMID: 25691825; PMCID: PMC4322527.

[Vollmer, 24] Vollmer, A., Hartmann, S., Vollmer, M., Shavlokhova, V., Brands, R. C., Kübler, A., & Saravi, B. (2024). Multimodal artificial intelligence-based pathogenomics improves survival prediction in oral squamous cell carcinoma. Scientific Reports, 14(1), 5687.

[Wang, 23] Wang, Q., Zhao, Y., Wang, F., & Tan, G. (2023). Clustering and machine learning-based integration identify cancer associated fibroblasts genes' signature in head and neck squamous cell carcinoma. Frontiers in Genetics, 14, 1111816.

[Wang, 21] Wang, X., & Li, B. B. (2021). Deep learning in head and neck tumor multiomics diagnosis and analysis: review of the literature. Frontiers in Genetics, 12, 624820.

[Wang, 23] Wang, Y., He, Y., Duan, X., Pang, H., & Zhou, P. (2023). Construction of diagnostic and prognostic models based on gene signatures of nasopharyngeal carcinoma by machine learning methods. Translational Cancer Research, 12(5), 1254.

[Way, 18] Way, G. P., & Greene, C. S. (2018). Extracting a biologically relevant latent space from cancer transcriptomes with variational autoencoders. In PACIFIC SYMPOSIUM on BIOCOMPUTING 2018: Proceedings of the Pacific Symposium (pp. 80-91).

[Weusthof, 23] Weusthof, C., Burkart, S., Semmelmayer, K., Stögbauer, F., Feng, B., Khorani, K., & Hess, J. (2023). Establishment of a Machine Learning Model for the Risk Assessment of Perineural Invasion in Head and Neck Squamous Cell Carcinoma. International Journal of Molecular Sciences, 24(10), 8938.

[Worsham, 13] Worsham, M. J., Chen, K. M., Meduri, V., et al. (2013). Epigenetic events of disease progression in head and neck squamous cell carcinoma. Otolaryngology–Head and Neck Surgery, 149(3), 409-416. https://doi.org/10.1177/0194599813494996

[Wu, 22] Wu, S., Lv, X., Li, Y., Gao, X., Ma, Z., Fu, X., & Li, Y. (2022). Integrated machine learning and single-sample gene set enrichment analysis identifies a TGF-beta signaling pathway derived score in headneck squamous cell carcinoma. Journal of Oncology, 2022.

[Xie, 20] Xie, F., Zhang, J., Wang, J., Reuben, A., Xu, W., Yi, X., & Xia, T. (2020). Multifactorial deep learning reveals pan-cancer genomic tumor clusters with distinct immunogenomic landscape and response to immunotherapy. Clinical Cancer Research, 26(12), 2908-2920.

[Yates, 20] Yates, A. D., Achuthan, P., Akanni, W., Allen, J., Allen, J., Alvarez-Jarreta, J., & Flicek, P. (2020). Ensembl 2020. Nucleic acids research, 48(D1), D682-D688.

[Zhao, 20] Zhao, Z., Li, Y., Wu, Y., & Chen, R. (2020). Deep learning-based model for predicting progression in patients with head and neck squamous cell carcinoma. Cancer Biomarkers, 27(1), 19-28.