# A Comparative Study of CNN for Prediction of Human Cancer Types Integrating Protein-Protein Interaction Networks and Omics Data

**Marilio Freire de Almeida**[1], **Sérgio Nery Simões**[1], **Karin Satie Komati**[1]

[1]PPCOMP – Instituto Federal do Espírito Santo – Brazil

mariliofreire@gmail.com, {sergio, kkomati}@ifes.edu.br

***Abstract.*** *This paper investigates convolutional neural networks (CNN) for predicting cancer types by integrating protein-protein interaction (PPI) networks with omics data. While [Chuang et al. 2021] employed a single 3-layer CNN, we explore ten different architectures, including a custom model developed by our team (CNN2Layers), following their methodology. By evaluating the strengths and weaknesses of these models, we aim to identify the most effective CNN for accurately predicting various human cancers. Our proposed model achieved state-of-the-art performance using fewer layers. Interestingly, the simpler architectures achieved superior results, indicating their effectiveness in handling the specific characteristics of the dataset.*

## 1. Introduction

Cancer continues to be the second leading cause of death globally, with around 9.7 million fatalities and nearly 20 million new diagnoses estimated in 2022 [Bray et al. 2024]. Late-stage diagnoses are closely linked to lower survival rates, as advanced-stage patients have limited treatment options and worse prognoses [Allemani et al. 2018]. Additionally, tumor heterogeneity, encompassing genetic, epigenetic, and phenotypic variations, further complicates treatment by causing diverse responses and resistances to therapies [Dagogo-Jack and Shaw 2018].

In cancer prediction, integrating PPI networks with omics data enhances the understanding of disease mechanisms [Matsubara et al. 2019]. A PPI network maps the interactions between proteins within a cell, revealing how proteins collaborate in various biological processes [Barabási et al. 2010]. Omics data is a broader term encompassing various types of biological data at a molecular level, including genomics (study of the genome), transcriptomics (study of RNA transcripts, including gene expression profiles), proteomics (study of the proteome, or the full set of proteins), metabolomics (study of metabolites), and other related fields [Horgon and Kenny 2011]. Gene expression profiling is a subset of transcriptomics, which in turn is part of the broader field of omics and in cancer prediction, analyzing gene expression profiles can identify biomarkers and molecular signatures associated with different cancer types, aiding in early diagnosis and personalized treatment strategies [Qian et al. 2021].

Our proposal involves conducting a comparative study of different CNN models, inspired by the research detailed in the article by [Chuang et al. 2021]. We adopt the same methodology for databases and metrics as outlined in the reference article. That methodology includes generating 2D images from integrated PPI datasets, BioGRID, DIP, InAct,

MINH, and MIPS, and TCGA Database (The Cancer Genome Atlas), which provides extensive omics data, including gene expression profiles for various cancer types. Those images are provided as input to CNN models to recognize patterns and features within visual data. By training the CNN with those images, the model learns to predict cancer types based on the complex relationships and variations captured in the PPI and omics datasets.

In the study by [Chuang et al. 2021], the authors experimented with only one architecture (a 3-layer CNN). In contrast, this paper explores several CNN architectures, including LeNet-5, AlexNet, VGG16, VGG19, InceptionV3, ResNet50, Xception, and DenseNet, to conduct experiments and comparisons following the methodology outlined by [Chuang et al. 2021]. Additionally, we also investigated the performance of a custom CNN architecture developed by our team (referred to as CNN2Layers). By examining the strengths and weaknesses of each model, we aim to identify the most effective CNN models for accurately predicting different types of human cancers. Our proposed model achieved state-of-the-art performance with fewer layers.

The article is structured as follows: Section 2 presents related works, Section 3 describes the database and the CNN models, and Section 4 shows the experimental results. Finally, Section 5 presents the main conclusions of this paper.

## 2. Related Works

Recent advancements in the application of deep learning techniques, particularly CNN, have significantly enhanced cancer diagnosis, therapy, and prognostic evaluation in biomedical applications. For instance, [Zhou et al. 2020] utilized deep learning models, specifically the Inception V3 CNN architecture, to predict clinically negative axillary lymph node metastasis from ultrasound images in patients with primary breast cancer. The model was trained and validated using a dataset of ultrasound images of breast cancer patients, achieving superior accuracy and sensitivity when compared to radiologists, with the model reaching an accuracy rate of 94.6% and a sensitivity of 92.1%. This underscores the potential of CNN in achieving higher diagnostic accuracy and improving patient outcomes in breast cancer care.

In another study, [Lombardo et al. 2021] applied CNN to analyze time-to-event data for distant metastases in head and neck cancer patients. The model was trained on imaging data to predict the time until distant metastasis occurred. Utilizing imaging datasets from independent cohorts of head and neck cancer patients, the CNN model showed a reliability score of 87.5% in predicting distant metastasis, facilitating better patient stratification and personalized treatment. This study highlighted the potential utility of CNN in clinical settings, enabling more accurate prognostic evaluations and aiding in personalized treatment strategies for cancer patients.

Furthermore, [Ma and Zhang 2018] developed the OmicsMapNet tool, which employs deep learning to transform high-dimensional omics data into 2D images for cancer grade predictions. Using RNA-Seq expression data of diffuse gliomas, the data were reorganized into treemaps, capturing the hierarchical structure of genes. CNN trained on these images demonstrated high accuracy in classifying Low-Grade Gliomas (LGG) and Glioblastoma Multiforme (GBM), achieving an overall accuracy of 93.8%. The OmicsMapNet tool effectively utilized the hierarchical structure of genes to enhance cancer

grade predictions, showcasing a novel approach to leveraging high-dimensional biological data in cancer diagnostics.

In another notable study, [Matsubara et al. 2019] integrated PPI networks with gene expression profiles to develop a CNN model for lung cancer classification. The study gathered PPI data from databases such as BioGRID, DIP, IntAct, MINT, and MIPS, and combined this data with RNA-Seq gene expression profiles. Spectral clustering techniques were used to generate 2D representations of the PPI networks, which were then used to train the CNN model. The integration of diverse biological datasets improved the model's robustness and predictive accuracy, effectively classifying lung cancer subtypes with an accuracy of 91.2%. This research highlighted the importance of combining PPI networks with gene expression data to enhance the predictive capabilities of CNN models in cancer classification. These advances demonstrate the significant potential of CNN and other deep learning techniques to improve cancer diagnosis, stratification, and outcome prediction. By integrating multidimensional data and complex biological interrelations, these approaches offer new opportunities to enhance therapeutic management and increase patient survival rates.

## 3. Materials and methods

This study aims to compare the 3-layer CNN architecture proposed by [Chuang et al. 2021] with several alternative CNN architectures. Additionally, we developed a custom model, CNN2Layers, inspired by their methodology, to evaluate the performance and robustness of different approaches. The methodology proposed by [Chuang et al. 2021] serves as the foundation for predicting and classifying normal tissues and tumors across 11 cancer types. Initially, protein-protein interaction (PPI) and gene expression data were compiled to create two-dimensional visual representations using clustering algorithms. Clinical and RNA-Seq data encompassing 6,136 samples from 11 cancer types, sourced from TCGA, were analyzed alongside PPI data for 16,433 human proteins. The Laplacian method was employed to project the PPI network into a two-dimensional space, facilitating its integration with gene expression data. These images were provided as input to CNN model to discriminate between normal tissues and cancer types. Of these, 1,228 images (307 normal tissues and 921 tumors) were designated for training, while the remaining 4,908 images were used for validation. At this point, we extended the original methodology by incorporating nine additional CNN models alongside the initial architecture, enabling a comparative analysis across a total of ten models. This extension allowed for a thorough evaluation of model performance and robustness across varied architectures. Finally, metrics and boxplots were computed to evaluate the models' performance. The source codes and datasets were uploaded on GitHub[1]. We utilized the same dataset available in this repository[2].

### 3.1. Datasets

We adopted the same data integration method used in [Chuang et al. 2021] cancer classification research, which combines level 3 RNA-Seq data with clinical data from TCGA. The Figure 1 illustrates the integrated dataset, which consists of gene expression profiles

[1] https://github.com/marilioalmeida/cnn-cancer-prediction.git

[2] https://github.com/bioxgem/CNN_model/tree/main/Convolutional-neural-network-for-human-cancer-types-prediction

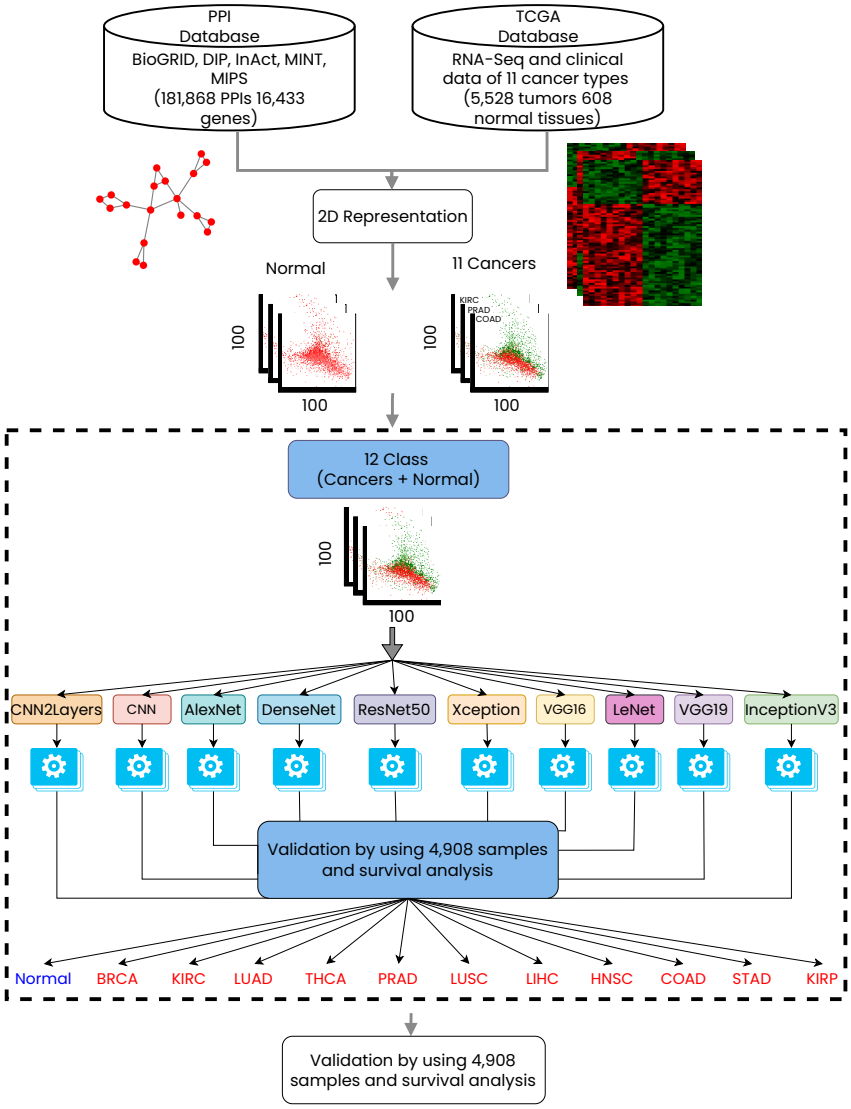


**Figure 1. Schematic overview of our method based on [Chuang et al. 2021]. Integration of PPI and RNA-Seq data from TCGA and PPI databases. The analysis encompasses 181,868 interactions involving 16,433 genes, alongside clinical data from 5,528 tumor samples and 608 normal tissue samples. In contrast to the approach of [Chuang et al. 2021], which employed a single architecture, our method applies 10 distinct CNN architectures to the same dataset (dotted box).**

for 20,531 genes derived from 5,528 tumor samples and 608 normal tissue samples. Table 1 presents a summary of the 11 distinct cancer types included in the integrated dataset, detailing their abbreviations, cancer type names, and the respective number of normal and tumor samples for each type. For the PPI integration, we followed the same methodology as [Chuang et al. 2021], where five public PPI databases were integrated: BioGRID [Oughtred et al. 2021], DIP [Xenarios et al. 2000], IntAct [Aranda et al. 2010], MINT [Calderone et al. 2020], and MIPS [Mewes et al. 2006], encompassing a total of 16,433 human proteins and 181,868 PPIs. Protein interactions were linked to gene expression profiles through gene names and IDs, creating a dataset with 14,230 proteins and 152,519 PPIs. Differentially expressed genes (DEGs) were determined using fold change and

modified t-statistic approaches, identifying 12,024 DEGs with notable expression changes in at least one cancer type. Spectral clustering was then applied to visualize these networks in two dimensions.

**Table 1. Expression datasets of RNA-seq and clinical data in 11 cancers.**

| Abbreviation | Cancer type | No. of normal | No. of tumor |
|---|---|---|---|
| BRCA | Breast invasive carcinoma | 113 | 1095 |
| COAD | Colon adenocarcinoma | 41 | 285 |
| HNSC | Head and neck squamous cell carcinoma | 44 | 520 |
| KIRC | Kidney renal clear cell carcinoma | 72 | 533 |
| KIRP | Kidney renal papillary cell carcinoma | 32 | 290 |
| LIHC | Liver hepatocellular carcinoma | 50 | 371 |
| LUAD | Lung adenocarcinoma | 59 | 515 |
| LUSC | Lung squamous cell carcinoma | 51 | 502 |
| PRAD | Prostate adenocarcinoma | 52 | 497 |
| STAD | Stomach adenocarcinoma | 35 | 415 |
| THCA | Thyroid carcinoma | 59 | 505 |

### 3.2. 2D Representation

For the 2D representation, we employed the same approach as described by [Chuang et al. 2021]. Specifically, they employed spectral clustering, utilizing the Laplacian (L) matrix to reduce the dimensionality of complex cancer networks, followed by the application of CNN techniques for further analysis. The cancer network is represented using an adjacency matrix $A$ and a diagonal matrix $D$, from which the Laplacian matrix $L$ is derived using the formula: $L = D - A$.

The symmetric Laplacian matrix (L) is structured such that cells are given a value of "-1" for connected protein pairs, while unconnected pairs are represented by "0". The diagonal cells contain the node degree, which corresponds to the number of edges each node has in the network. Subsequently, the eigenvalues and eigenvectors of the Laplacian matrix were derived through linear transformation. The researchers preserved the network's topology and connectivity by utilizing the smallest and second smallest non-zero, non-negative eigenvalues, along with their corresponding eigenvectors, to project the cancer network, consisting of 6,261 differently expressed genes and 28,439 interactions, onto a 2D grid with dimensions of 100×100 cells.

This technique projected the PPI network onto a two-dimensional plane, using the smallest eigenvalues for the x and y coordinates. The resulting network included 1,849 unique nodes, each representing a protein, with gene expression data from clinical samples assigned to corresponding nodes. In cases where multiple genes were mapped to a single node, their expression levels were averaged. A total of 6,136 cancer network images were created as input for a CNN model, which was tasked with differentiating between normal tissues, tumor samples, and specific cancer types.

### 3.3. Convolutional Neural Network

Convolutional Neural Networks, also known as ConvNets, are a class of deep learning models specifically designed for processing structured grid data, such as images [LeCun et al. 1989]. In [Chuang et al. 2021], the authors implemented a CNN architecture consisting of three consecutive convolutional layers, each containing 64 filters, referred to in this work as (CNN). The first convolutional layer uses a 5x5 kernel, while the

subsequent layers use 3x3 kernels, all with ‘same’ padding to maintain the spatial dimensions of the input. Batch normalization is applied after each convolutional layer to enhance the training process and model performance. Each convolutional layer is followed by a max-pooling layer with a 2x2 pool size and ’same’ padding to reduce the spatial dimensions. The network is then flattened and passed through four fully connected layers with 1000, 600, 80, and 12 neurons, respectively. The first three dense layers use ReLU activation, and the final layer uses softmax activation for classification.

We also modified the AlexNet CNN architecture by reducing three layers to simplify the network, resulting in a new model Figure 2 (CNN2layers) , the first convolutional layer uses a 11x11 kernel with a stride of 4x4 and ’valid’ padding, while the second convolutional layer uses a 5x5 kernel with a stride of 1x1 and ’same’ padding to maintain the spatial dimensions. Each convolutional layer is followed by a max-pooling layer with a 3x3 pool size and ’valid’ padding to reduce the spatial dimensions. The network is then flattened and passed through four fully connected layers with 4096, 4096, 1000, and 12 neurons, respectively. The first three dense layers use ReLU activation, and the final layer uses softmax activation for classification.

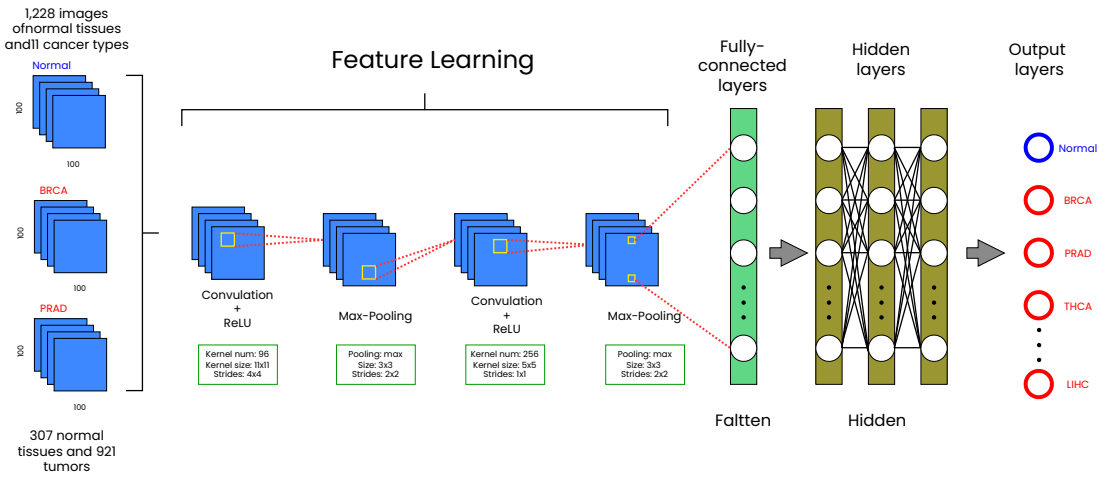


**Figure 2. The CNN2Layers architecture, which we implemented with two convolutional layers, was developed for classifying normal and tumor tissue images. The network processes a dataset of 1,228 images, including 307 normal tissue samples and 921 tumor samples from 11 distinct cancer types. The first convolutional layer uses 96 filters (11x11), while the second applies 256 filters (5x5), both utilizing ReLU activation functions. Each convolutional layer is followed by a 3x3 pooling layer. The extracted features are then passed through fully connected layers, culminating in the classification into 12 categories, encompassing both normal and cancerous tissues.**

Figure 3 illustrates the timeline of CNN model developments. Below, we provide a timeline overview of each CNN architecture utilized in our experiments.

- **LeNet-5** [LeCun et al. 1998] is a pioneering CNN architecture designed to recognize handwritten digits. The LeNet-5 architecture comprises three convolutional layers, two sub-sampling (pooling) layers, and two fully connected layers. It utilizes backpropagation for training and contains a total of 431,000 weights.

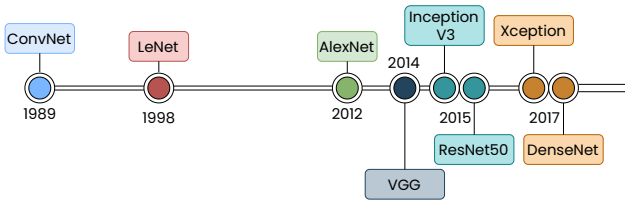


**Figure 3. Timeline of CNN models.**

- **AlexNet** [Krizhevsky et al. 2012] represented a significant advancement over its predecessor, LeNet. This CNN model was designed to classify 1,000 image classes from the ImageNet LSVRC-2010 dataset, achieving groundbreaking results at the time. AlexNet consists of eight layers: five convolutional layers and three fully connected layers. It features 60 million parameters and 650,000 neurons, highlighting its complexity and capacity for DL.
- **VGG16 and VGG19** [Simonyan and Zisserman 2014] introduced as one of the most popular submissions to ILSVRC-2014, the VGG models, including VGG-11, VGG-16, and VGG-19, consist of 11, 16, and 19 layers, respectively. The primary objective of varying the number of convolutional layers (Conv2D) was to better understand the impact of network depth on image classification accuracy. VGG-19, the most computationally intensive model, contains 138 million weights.
- **InceptionV3** [Szegedy et al. 2015] is the third iteration of the Inception architecture and incorporates several innovative techniques, such as factorized convolutions, aggressive regularization, and auxiliary classifiers, to improve both efficiency and accuracy. InceptionV3 achieves high performance by using smaller convolutions to reduce the number of parameters and computational complexity while maintaining high representational power.
- **ResNet50** (Residual Network) [He et al. 2016] achieved the first place in the 2015 ILSVRC challenge. The primary goal of ResNet was to address the vanishing gradient problem that hindered previous deep networks. The ResNet-50 architecture comprises 49 convolutional layers followed by one fully connected layer, totaling 25.5 million weights.
- **Xception** [Chollet 2017] builds upon the inception model by replacing the standard Inception modules with depthwise separable convolutions. This modification allows Xception to achieve greater efficiency and performance by independently processing spatial and channel-wise information within the network. By leveraging the strengths of depthwise separable convolutions, Xception effectively reduces the number of parameters and computational cost.
- **DenseNet** [Huang et al. 2017] features a unique architecture where each layer is densely connected to every other layer. This interconnected structure is what gives DenseNet its name. Each layer receives input from all preceding layers and passes its feature maps to all subsequent layers, promoting extensive feature reuse. The DenseNet architecture comprises multiple dense blocks, with transition blocks in between to manage dimensionality and ensure efficient processing.

In this study, we employed several CNN models to perform experiments and comparisons, following the methodology proposed by [Chuang et al. 2021]. The goal was to train these networks to distinguish between normal conditions and 11 different types of

cancer. The main contribution of this work is demonstrating that by modifying and simplifying the CNN2Layers model, we achieved results comparable to, or exceeding, state-of-the-art performance with fewer layers, while preserving high accuracy and robustness. This simplification significantly reduced computational complexity without sacrificing performance, representing a notable advancement in optimizing CNN architectures for cancer classification.

## 4. Experiments and Results

In Section 3, we presented the methodology proposed by the researchers [Chuang et al. 2021], which integrates PPI and TCGA datasets to generate 6,136 images used as input for a CNN network. In this study, we applied this dataset as input to 10 different CNN architectures for comparative analysis, including the architecture employed by the aforementioned researchers. The goal was to train the various architectures to classify between normal conditions and 11 different types of cancer. Out of the 10 architectures evaluated, one (a classical CNN) was the same as used in the work by [Chuang et al. 2021], while the remaining eight were state-of-the-art models. Additionally, we modified the AlexNet CNN architecture by reducing three layers to simplify the network, resulting in a new model (CNN2layers).

To maintain consistency and comparability, we adopted the same data partitioning strategy used by [Chuang et al. 2021], with 1,228 images (307 normal tissues and 921 tumors) for training, and the remaining 4,908 images for validation. Each model was trained across 50 rounds, with 100 epochs per round, and evaluated using the metrics of Accuracy, Precision, Recall, F1-Score, and ROC AUC Score.

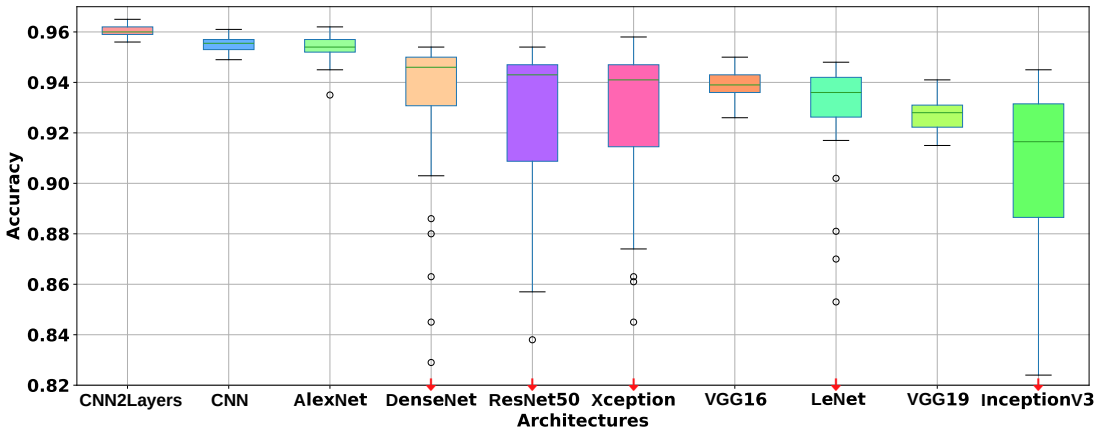


**Figure 4. Accuracy comparison across various neural network architectures. The red arrows indicate more outliers below. The figure demonstrates that the CNN2Layers architecture achieved the highest accuracy and exhibited the lowest variance, indicating strong robustness.**

Figure 4 presents the accuracy results for various neural network architectures. The X-axis represents the architectures compared—CNN2Layers, CNN, AlexNet, DenseNet, ResNet50, Xception, VGG16, LeNet, VGG19, and InceptionV3—while the Y-axis shows accuracy values, ranging from approximately 0.82 to 0.96. The architecture with the highest accuracy was CNN2Layers, achieving 0.96, followed closely by CNN (0.9555) and AlexNet (0.954). These models demonstrated exceptional performance,

yielding the top accuracy scores among the tested architectures. In terms of variance, CNN2Layers, CNN, and AlexNet exhibited low variance, suggesting they are robust architectures with consistent performance across different runs or datasets. Conversely, InceptionV3, while still performing well with an accuracy of 0.9165, showed lower accuracy compared to the leading models. Notably, while DenseNet achieved a relatively high accuracy of 0.946, it displayed greater variance, potentially indicating sensitivity to data variations or less consistent performance. ResNet50 and Xception also yielded satisfactory results, with accuracies of 0.943 and 0.941, respectively, though they were slightly outperformed by the top models. Lastly, VGG19 and VGG16 recorded accuracies of 0.928 and 0.939, positioning them mid-tier in terms of performance, with similar results, suggesting comparable behavior. Despite being an older architecture, LeNet remained competitive with an accuracy of 0.936, underscoring its effectiveness even in light of more recent advancements.

The Figure 5 illustrates the F1 Score results for various neural network architectures, comparing their performance based on a balance between precision and recall. The X-axis represents the architectures evaluated, while the Y-axis denotes the F1 Score values, ranging from 0.82 to 0.96. Among the architectures, CNN2Layers achieved the highest F1 Score, reaching 0.96, indicating superior performance overall. Similarly, CNN also performed well, with an F1 Score of 0.9555, suggesting that both architectures are strong candidates for this task. Conversely, the InceptionV3 architecture recorded the lowest F1 Score, at 0.9165, highlighting its comparatively weaker performance. A noteworthy observation is the low variance in performance for certain architectures, such as CNN2Layers, CNN, and AlexNet, indicating greater robustness and stability, with minimal outliers. This stability is crucial for ensuring consistent performance across different application scenarios. In conclusion, the results clearly demonstrate the superior performance of CNN2Layers, with CNN and AlexNet closely following, while InceptionV3 exhibited the most modest performance among the evaluated architectures.

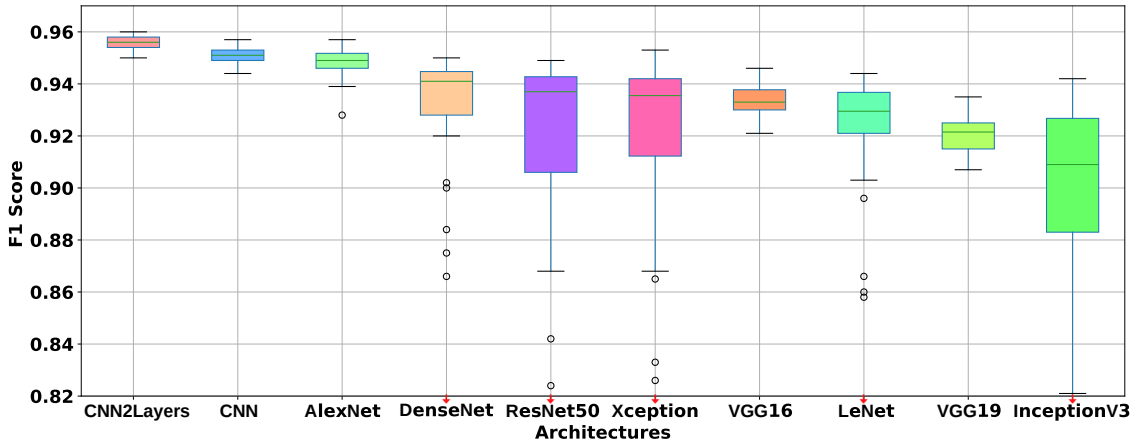


**Figure 5. F1-Score comparison across various neural network architectures. The red arrows indicate more outliers below. The figure demonstrates that the CNN2Layers architecture achieved the highest F1-score and also exhibited the lowest variance, indicating strong robustness.**

The Table 2 summarizes the performance of various CNN architectures, evalu-

| Architectures | Accuracy | Precision | Recall | F1-Score | ROC AUC Score |
|---|---|---|---|---|---|
| **CNN2Layers** | **0.960** | **0.987** | **0.958** | **0.956** | 0.998 |
| CNN | 0.9555 | 0.985 | 0.953 | 0.951 | 0.998 |
| AlexNet | 0.954 | 0.982 | 0.9515 | 0.949 | 0.998 |
| DenseNet | 0.946 | 0.978 | 0.942 | 0.941 | 0.996 |
| ResNet50 | 0.943 | 0.973 | 0.9395 | 0.937 | 0.995 |
| Xception | 0.941 | 0.974 | 0.9375 | 0.9355 | 0.995 |
| VGG16 | 0.939 | 0.969 | 0.9365 | 0.933 | 0.996 |
| LeNet | 0.936 | 0.981 | 0.933 | 0.9295 | 0.998 |
| VGG19 | 0.928 | 0.959 | 0.924 | 0.9215 | 0.995 |
| InceptionV3 | 0.9165 | 0.963 | 0.9075 | 0.909 | 0.9925 |

**Table 2. Median values of Accuracy, Precision, Recall, F1-Score, and ROC AUC Score for different architectures, ordered by Accuracy.**

ated using metrics such as Accuracy, Precision, Recall, F1-Score, and ROC AUC Score. Among these, the CNN2Layers model emerges as the top performer across multiple metrics. It achieved the highest accuracy (0.960), indicating the most correct predictions relative to the total number of examples, and the highest precision (0.987), demonstrating a strong ability to correctly predict positive classes. Additionally, CNN2Layers led in recall (0.958), meaning it effectively identified the majority of positive instances, and recorded the best F1-Score (0.956), signifying an optimal balance between precision and recall. Furthermore, several models, including CNN2Layers, CNN, AlexNet, and LeNet, attained the highest ROC AUC Score (0.998), reflecting excellent class discrimination capabilities.

Consequently, CNN2Layers is the most balanced model overall, making it a suitable choice for applications that require both high accuracy and reliable identification of positive instances. While LeNet did not achieve the highest accuracy or precision, its ROC AUC Score of 0.998 aligns it with more advanced models like CNN2Layers and AlexNet, underscoring its strong class separation ability. In contrast, InceptionV3, with an accuracy of 0.9165, ranks among the lower-performing models in this evaluation. However, its relatively high precision (0.963) may be advantageous in contexts where minimizing false positives is more critical than capturing all true positives.

## 5. Conclusion

This study highlights the potential of CNN for cancer type prediction and classification by integrating PPI networks with omics data. By leveraging a variety of CNN architectures (CNN2Layers, CNN, AlexNet, DenseNet, ResNet50, Xception, VGG16, LeNet, VGG19, and InceptionV3) we were able to assess and compare the performance, accuracy, and computational efficiency of each model. Interestingly, simpler models demonstrated superior results.

The integration of PPI networks with gene expression data from the TCGA database allowed CNN to capture complex biological interactions, leading to improved predictive accuracy. Unlike [Chuang et al. 2021] study, which employed a single CNN architecture, our analysis encompassed 10 distinct architectures, offering a broader evaluation. Notably, we achieved state-of-the-art performance using fewer layers, showcasing both enhanced computational efficiency and predictive capability. These findings underscore the promise of CNN models, particularly when enriched with comprehensive biological data, in advancing personalized cancer diagnosis and treatment strategies.

Future research could explore data augmentation techniques to improve model robustness, particularly for handling imbalanced cancer type distributions. Additionally, testing the models on external datasets beyond the TCGA is necessary to evaluate their generalizability. Finally, applying these models in clinical settings would be critical for assessing their practical impact on patient care outcomes.

## Acknowledgment

The authors thank IFES, and CAPES/FAPES for the PDPG (Graduate Development Program, Strategic Partnerships in the States) – SIAFEM nº 2021-2S6CD, nº FAPES 132/2021. Professor Komati thanks CNPq for the DT-2 grant (nº 302726/2023-3) and project nºº407742/2022-0, also thanks FAPES (Fundação de Amparo à Pesquisa e Inovação do Espírito Santo) for project nº 1023/2022 P:2022-8TZV6.

## References

Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Nikšić, M., Bonaventure, A., Valkov, M., Johnson, C. J., Estève, J., et al. (2018). Global surveillance of trends in cancer survival 2000–14 (concord-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *The Lancet*, 391(10125):1023–1075.

Aranda, B., Achuthan, P., Alam-Faruque, Y., Armean, I., Bridge, A., Derow, C., Feuermann, M., Ghanbarian, A., Kerrien, S., Khadake, J., et al. (2010). The intact molecular interaction database in 2010. *Nucleic acids research*, 38(suppl_1):D525–D531.

Barabási, A., Gulbahce, N., and Loscalzo, J. (2010). Network medicine: a network-based approach to human disease. *Nature Reviews Genetics*, 12:56–68.

Bray, F., Laversanne, M., Sung, H., Ferlay, J., Siegel, R. L., Soerjomataram, I., and Jemal, A. (2024). Global cancer statistics 2022: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 74(3):229–263.

Calderone, A., Iannuccelli, M., Peluso, D., and Licata, L. (2020). Using the mint database to search protein interactions. *Current Protocols in Bioinformatics*, 69(1):e93.

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1251–1258.

Chuang, Y.-H., Huang, S.-H., Hung, T.-M., Lin, X.-Y., Lee, J.-Y., Lai, W.-S., and Yang, J.-M. (2021). Convolutional neural network for human cancer types prediction by integrating protein interaction networks and omics data. *Scientific reports*, 11(1):20691.

Dagogo-Jack, I. and Shaw, A. T. (2018). Tumour heterogeneity and resistance to cancer therapies. *Nature reviews Clinical oncology*, 15(2):81–94.

He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778.

Horgon, R. and Kenny, L. (2011). Sac review: Omic technologies: genomics, transcriptomics. *Proteomics and metabolomics. TOG*, 13(189-195):25.

Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.

Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.

LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., and Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, 1(4):541–551.

LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324.

Lombardo, E., Kurz, C., Marschner, S., Avanzo, M., Gagliardi, V., Fanetti, G., Franchin, G., Stancanello, J., Corradini, S., Niyazi, M., et al. (2021). Distant metastasis time to event analysis with cnns in independent head and neck cancer cohorts. *Scientific reports*, 11(1):6418.

Ma, S. and Zhang, Z. (2018). Omicsmapnet: Transforming omics data to take advantage of deep convolutional neural network for discovery. *arXiv preprint arXiv:1804.05283*.

Matsubara, T., Ochiai, T., Hayashida, M., Akutsu, T., and Nacher, J. C. (2019). Convolutional neural network approach to lung cancer classification integrating protein interaction network and gene expression profiles. *Journal of Bioinformatics and Computational Biology*, 17(03):1940007. PMID: 31288636.

Mewes, H.-W., Frishman, D., Mayer, K. F., Münsterkötter, M., Noubibou, O., Pagel, P., Rattei, T., Oesterheld, M., Ruepp, A., and Stümpflen, V. (2006). Mips: analysis and annotation of proteins from whole genomes in 2005. *Nucleic acids research*, 34(suppl_1):D169–D172.

Oughtred, R., Rust, J., Chang, C., Breitkreutz, B.-J., Stark, C., Willems, A., Boucher, L., Leung, G., Kolas, N., Zhang, F., et al. (2021). The biogrid database: A comprehensive biomedical resource of curated protein, genetic, and chemical interactions. *Protein Science*, 30(1):187–200.

Qian, Y., Daza, J., Itzel, T., Betge, J., Zhan, T., Marmé, F., and Teufel, A. (2021). Prognostic cancer gene expression signatures: Current status and challenges. *Cells*, 10(3).

Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., and Rabinovich, A. (2015). Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9.

Xenarios, I., Rice, D. W., Salwinski, L., Baron, M. K., Marcotte, E. M., and Eisenberg, D. (2000). Dip: the database of interacting proteins. *Nucleic acids research*, 28(1):289–291.

Zhou, L.-Q., Wu, X.-L., Huang, S.-Y., Wu, G.-G., Ye, H.-R., Wei, Q., Bao, L.-Y., Deng, Y.-B., Li, X.-R., Cui, X.-W., et al. (2020). Lymph node metastasis prediction from primary breast cancer us images using deep learning. *Radiology*, 294(1):19–28.