

Article

Estimating the Prognosis of Low-Grade Glioma with Gene Attention Using Multi-Omics and Multi-Modal Schemes

Sanghyuk Roy Choi  and Minhyeok Lee * 

School of Electrical and Electronics Engineering, Chung-Ang University, Seoul 06974, Korea

* Correspondence: mlee@cau.ac.kr

Simple Summary: The estimation of the prognosis of low-grade glioma (LGG) patients using deep learning models and gene expression data has been intensively studied in recent years. Existing studies, however, have only considered mRNA expression data, ignoring other expression data and clinical data. The Multi-Prognosis Estimation Network (Multi-PEN), a deep learning model that employs multi-omics and multi-modal schemes, is proposed in this study to address this limitation. Using Multi-PEN, MYBL1 and hsa-mir-421 were identified as the most significant mRNA and miRNA, respectively, in the prognosis of LGG patients. Existing studies that estimate prognostic mRNAs and miRNAs support the findings of this study.

Abstract: The prognosis estimation of low-grade glioma (LGG) patients with deep learning models using gene expression data has been extensively studied in recent years. However, the deep learning models used in these studies do not utilize the latest deep learning techniques, such as residual learning and ensemble learning. To address this limitation, in this study, a deep learning model using multi-omics and multi-modal schemes, namely the Multi-Prognosis Estimation Network (Multi-PEN), is proposed. When using Multi-PEN, gene attention layers are employed for each datatype, including mRNA and miRNA, thereby allowing us to identify prognostic genes. Additionally, recent developments in deep learning, such as residual learning and layer normalization, are utilized. As a result, Multi-PEN demonstrates competitive performance compared to conventional models for prognosis estimation. Furthermore, the most significant prognostic mRNA and miRNA were identified using the attention layers in Multi-PEN. For instance, MYBL1 was identified as the most significant prognostic mRNA. Such a result accords with the findings in existing studies that have demonstrated that MYBL1 regulates cell survival, proliferation, and differentiation. Additionally, hsa-mir-421 was identified as the most significant prognostic miRNA, and it has been extensively reported that hsa-mir-421 is highly associated with various cancers. These results indicate that the estimations of Multi-PEN are valid and reliable and showcase Multi-PEN's capacity to present hypotheses regarding prognostic mRNAs and miRNAs.

Keywords: prognosis estimation; low-grade glioma; deep learning; gene expression; prognostic gene; ensemble learning; attention mechanism



Citation: Choi, S.R.; Lee, M. Estimating the Prognosis of Low-Grade Glioma with Gene Attention Using Multi-Omics and Multi-Modal Schemes. *Biology* **2022**, *11*, 1462. <https://doi.org/10.3390/biology11101462>

Academic Editors: Chung-Der Hsiao and Tzong-Rong Ger

Received: 31 August 2022

Accepted: 2 October 2022

Published: 5 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Owing to the rapid development of deep learning models, artificial intelligence based on deep learning has dominated various fields [1–3]. Deep learning models are being used extensively in diagnostics [4–6], medical imaging [7–10], and genome sequencing [11–13]. Additionally, the prognosis of various diseases can be estimated by deep learning models [14–16]. In these models, clinical variables, such as age and sex, are often utilized to predict prognosis. This study aims to propose a deep learning model for the estimation of the prognosis of low-grade glioma (LGG) [17]. Tumors in the central nervous system are defined by their origin and their histopathological features [18]. Gliomas are neuroepithelial tumors stemming from the supporting glial cells located in the central nervous system.

LGG is a group of primary brain tumors frequently generated in young and healthy patients. LGG consists of grade one tumors and grade two tumors. Grade I LGG tumors do not involve histologic features such as atypia, mitotic activity, or necrosis. In contrast, grade II LGG tumors do involve the aforementioned histologic features [19]. Symptoms of LGG include headaches, vomiting, blurry vision, memory loss, nausea, and weakness on one side of the body [20–23]. LGG causes seizures more frequently than other brain cancers because it is produced in the cerebral cortex. Generally, LGG patients show a better prognosis compared to high-grade glioma (HGG) patients. However, if it progresses to HGG, the prognosis may drastically deteriorate, eventually leading to death [24]. As a result, the early diagnosis and treatment of LGG are essential in order to increase the likelihood of survival.

Gene expression has become remarkably significant in the identification of tumors. For instance, mutations of isocitrate dehydrogenase 1 and 2 (IDH1 and IDH2) act as tumor markers [19]. Additionally, the capicua transcriptional repressor (CIC) gene on chromosome 19q has recently been extensively explored, where oligodendroglioma and oligoastrocytomas cooperate with the mutation of the CIC gene [25,26]. CIC gene mutations are discovered in 69% of oligodendrogliomas. Additionally, mutations of IDH1 and IDH2 and 1p/19q codeletion have been proven to collaborate with CIC gene mutations. Moreover, the overexpression and mutation of tumor protein 53 (TP53) are described as a genetic feature of gemistocytic astrocytoma and low-grade astrocytoma.

Therefore, the relationship between gene expression and prognosis in LGG has been widely studied to estimate the prognosis with deep learning models using gene expression. For instance, the Gene Attention Ensemble NETwork (GAENET) was proposed as a deep learning model to estimate the prognosis of LGG patients [27]. In GAENET, the gene attention mechanism, which is a modification of the attention mechanism [28–30], was introduced. The gene attention layer scores genes with respect to prognosis. The value of the scores ranges from zero to one, in which a high value indicates a close relationship with the prognosis. Since the gene attention layer is trained by a deep learning process, nonlinear relationships are studied to select the prognostic genes, which is difficult to do in other conventional models. Additionally, residual learning [31] and ensemble learning methods [32] are employed in GAENET, which uses deep learning methods to enhance the prediction performance.

Numerous studies have demonstrated that using multi-omics and multi-modal data results in improved performance [33–35]. For example, a study proposed two models using multi-omics data: a risk score model that estimated the prognosis of LGG patients based on six specific genes and a radiomic risk score model that exploited magnetic resonance images to predict the pre-operation risk score [33]. Another study used a multivariate Cox-PH model with various data. It applied numerous combinations of clinical, radiomics, and molecular data [34]. An integrated multi-omics deep learning network method (i-Modern) uses one autoencoder to extract significant elements from multi-omics data [35].

Therefore, a crucial limitation of GAENET is that a limited transcriptome data type is used in the model. Specifically, GAENET only uses mRNA data, but other transcriptome data types are expected to be related to the prognosis. For instance, it is widely known that miRNA controls mRNA by regulating mRNA and hindering its expression [36]. Such a property signifies that miRNA has the potential to be highly related to the prognosis of LGG patients. Additionally, it is generally accepted that clinical data are also highly related to prognosis.

This paper proposes a deep learning model that employs not only mRNA data but also miRNA and clinical data. The motivation for this study comes from the idea that such a perspective with multi-omics and multi-modal schemes can provide a better understanding of the prognosis of LGG patients. Given this additional information relating to miRNA, the identified prognostic genes may be distinct from those identified when using GAENET and may explain the relationship between mRNA and miRNA in terms of patient prognosis. While there are several studies that take advantage of the high correlation between mRNA

and miRNA [37,38], in this study, a multi-modal scheme [39] was additionally used with the clinical data. Moreover, via gene attention layers that use miRNA data and clinical data as inputs, it can be estimated which miRNA is most likely to be related to the prognosis of LGG.

2. Background

2.1. Gene Attention Layer

Deep learning models commonly consist of multiple layers. Each layer comprises an input, a weight matrix, and nonlinear activation functions. The layer receives input from the last layer and multiplies the input and weight matrix [40]. The values made from multiplying the input and output pass through the nonlinear activation function and are transferred to the subsequent layer. This process is repeatedly conducted until it reaches the output layer. This process can be represented by the following equation:

$$\hat{Y} = \tau_N (W_N \cdot \tau_{N-1} (W_{N-1} \cdots \tau_2 (W_2 \cdot \tau_1 (W_1 \cdot X))), \quad (1)$$

where \hat{Y} is an estimation vector that is the output of the deep learning model; τ_k and W_k are the nonlinear function and weight matrix in the k^{th} layer, respectively; X is an input vector of the deep learning model; and N is the total number of layers used in the model. Training a deep learning model is identical to finding an appropriate W_k to estimate \hat{Y} properly for a given \hat{X} .

The attention mechanism is a method of multiplying the output vectors of τ_k by specific values from zero to one. This method reduces the effects of features less related to the target. Thus, a lower value, approximately zero, indicates that the corresponding feature has a low probability of having a relationship with the target. This process can be considered a filter to eliminate less related features. The mechanism can be represented by the following equation:

$$X_{k+1} = A^T \odot \tau_k (W_k \cdot X_k), \quad (2)$$

where X_k is the output of the $(k - 1)^{\text{th}}$ intermediate layer; A is an attention vector composed of the attention values of all input features of X_k ; and the \odot operator denotes component-wise multiplication. The attention vector A is not a constant and is instead a deep learning module that learns and changes its value in every epoch based on the backpropagation in deep learning training.

The attention layer in GAENET uses a squeeze and excitation (SE) block composed of two fully connected networks. Thus, the block has two weight matrices, and the first weight matrix is multiplied by gene expression inputs; it reduces the dimension of the gene expression, resulting in feature extraction. Therefore, the first weight matrix has a dimension of $q \times p$, where p is the number of genes and $q \ll p$. The next layer, accordingly, has a dimension of $p \times q$, producing an attention vector with p dimension. These processes can be represented as follows:

$$A = \sigma(W_{A2} \cdot \text{ReLU}(W_{A1} \cdot X_0)), \quad (3)$$

$$X_1 = A^T \odot X_0, \quad (4)$$

where σ denotes the sigmoid activation function; W_{A1} and W_{A2} indicate the first and second weight matrices of the attention layer, respectively; $X_0 \in R^p$ is the gene expression input; and $X_1 \in R^p$ is the output of the attention layer. Consequently, X_1 becomes the input of the deep learning model following the attention layer.

Owing to the attention vector A with values between zero and one, the input gene expression data are regulated before they are transferred to the subsequent deep learning layers. The attention layer assigns values close to 1.0 to highly target-related genes, whereas it assigns values close to 0.0 to the genes considered less related to the target. Two weight matrices in the attention layer are trained during the training process simultaneously with

the other deep learning layers. This process shows that the attention value is determined via training on gene expression data and is not manual.

Therefore, the attention value of each gene tells us how important the gene is. Hence, the target-related genes can be discovered. In other words, the prognostic genes can be found by the attention layer when the prognosis is trained to be predicted by the model. Similar to other attention mechanisms, this gene attention layer also regulates deep learning models to avoid overfitting models, reducing overfitting problems caused by a lack of data.

2.2. Residual Learning

The vanishing gradient problem is a common problem for deep learning models that use many layers. The cause of the vanishing gradient problem is multiplied gradients in the chain rule. The backpropagation algorithm used to train deep learning models calculates the gradients of the initial layers via the chain rule, which involves multiplying multiple gradients. If the gradients are between zero and one, the multiplication of the gradients makes the gradients of the initial layers converge upon zero. Since the conventional training algorithm for deep learning models depends on the gradients, significantly low gradient values in the initial layers cause a training failure in the layers.

Residual learning with a residual network is a method to alleviate the vanishing gradient problem. Residual networks consist of a series of layers with a skip connection between the input of the first layer and the output of the last layer [31]. In this connection, the input is added to the output, resulting in a capability to simplify the multiplication in the chain rule. Since the vanishing gradient problem is caused by a series of multiplications in the chain rule, such a simplification can reduce the problem. Therefore, residual learning enables us to stack more layers compared to conventional deep learning models. Residual networks with two layers can be represented as follows:

$$X_{k+2} = X_k + \tau_{k+1} (W_{k+1} \cdot \tau_k (W_k \cdot X_k)), \quad (5)$$

where X_k is the output of the previous layer.

2.3. Ensemble Learning

Ensemble learning is a method to prevent deep learning models from overfitting by using multiple models on the same target [32,41]. Due to the fact that it has more parameters than samples in general, deep learning can result in overfitting issues. Such deep learning models with overfitting issues make accurate predictions from the training set, while the test set accuracy declines. In order to find a solution to the overfitting issue, the method of ensemble learning is utilized. In ensemble learning, each of a number of modules acquires new knowledge from the data. Then, the next step is to select the most effective modules to demonstrate the entire model or to use the result that represents a weighted sum of all of the modules as the final result. The model can also be protected from outliers, which can be removed by averaging the results of multiple learning. Moreover, ensemble learning can address problems caused by the random selection of inputs for mini-batch sampling. If outliers are included in the training set for a mini-batch, a single model has a probability of diverging or learning inappropriate features. However, ensemble learning methods effectively circumvent the aforementioned difficulties with the effect of multiple modules.

In this study, a deep learning ensemble model with multiple modules is presented. Each module in the ensemble deep learning model employs the same structure. Due to the random initialization of weight parameters and mini-batch sampling, even when the same structure is employed, these modules generate distinct results. These diverse results can counteract the randomness of conventional deep learning training, in which the results vary. As the results of multiple modules are averaged in the ensemble deep learning model, this issue can be mitigated. The final prediction derived from the ensemble deep learning model is computed as follows:

$$\hat{Y} = \frac{1}{n} \sum_{i \in} f_i(X) \quad (6)$$

where n is the number of modules in the ensemble deep learning and f_i represents a module with an equivalent structure.

2.4. Layer Normalization

Layer normalization is a regularization method to maintain the output distributions of the layers [42]. During the training process of deep learning, the output distributions consistently fluctuate. This property of training can hinder the convergence of training since it can be interpreted that each layer learns different sample spaces in terms of the input of the layer. Thus, the optimal convergence point also fluctuates, resulting in the possibility of failure in training. To address this limitation, layer normalization is adopted in the proposed model. In the residual learning modules, layer normalization is employed before activation functions, which is the conventional way to use layer normalization.

2.5. The Gene Attention Ensemble Network

The GAENET, which is one of the baselines of the proposed model, employs mRNA to predict the prognosis of LGG patients. To handle mRNA data and extract the features regarding the prognosis, the attention layers are used as the first layer. Then, two residual modules are trained in the ensemble structure. In these modules, layer normalization and dropout are used. As a result of the prognosis estimation, the average outcome of these modules is computed. However, in GAENET, mRNA data are employed, which can be one of the limitations since deep learning training with different data types can consider different aspects of the prognosis of LGG. This limitation motivated us to explore a novel method of using multi-omics and multi-modal schemes integrated with a deep learning model.

2.6. Baselines

Four conventional machine learning models to estimate survival times are employed as baselines: ridge regression, survival support vector machine (SurvivalSVM) [43], random survival forest (RSF) [44], and Coxnet [45]. Ridge regression is a method that can be utilized to analyze data that have multicollinearity or to estimate a prediction vector. This method exploits L2 regularization and can handle the problem in conventional linear regression wherein the prediction vectors in multiple regression are not orthogonal, so parameter estimation based on the minimum sum of squares is likely to be poor. Coxnet is a regularized Cox proportional hazard with L1, L2, and both L1 and L2. A random forest is a probabilistic predictor that uses a mean to improve prediction accuracy and drops over-fitting when fitting several decision tree classifiers. A support vector machine is one of the supervised machine learning models. Support vector machines are commonly used in classification tasks to reduce the dimensions of the dataset.

3. Methods

3.1. Multi-Omics and Multi-Modal Ensemble Deep Learning

Numerous functions of RNAs have been discovered and studied as a result of technological advances in genome sequencing. It has been widely accepted that mRNA is highly correlated with miRNA, which regulates the function of mRNA and exerts a significant influence on diseases. Consequently, the model proposed in this paper, the Multi-Prognosis Estimation Network (Multi-PEN), utilizes multi-omics data consisting of mRNA and miRNA [46,47]. Two attention layers are utilized for multi-omics data, each of which identifies the prognostic genes and assigns a high attention value to significant genes. As was mentioned in the previous section, the learning of attention values is also performed by the backpropagation algorithm of deep learning, which can be interpreted as learning from data.

Additionally, clinical data, such as age at diagnosis and sex, are known to be associated with the prognosis of LGG, as many studies have aimed to predict the prognosis using such variables. In Multi-PEN, the clinical data are employed as one of the multi-modal inputs of

the model. Since there are several categorical variables in the clinical data, they are encoded as one-hot vectors; since the computer machine is incapable of perceiving human language, clinical categorical variables must be encoded into vectors consisting of zero and one, of which an index of one indicates the status of the corresponding category, such as sex or features of tumors comprising clinical data. Then, all clinical variables are concatenated to a single vector. This clinical vector is also trained with an attention layer, the same as mRNA and miRNA datatypes.

Figure 1 depicts the structure of Multi-PEN. Three different datatypes with multi-omics and multi-modal schemes are employed as the inputs. These inputs are transferred to the attention layers and multiplied by the attention values that are learned with each data type. The attention values range from zero to one, where higher values indicate greater association with the target, i.e., the prognosis of LGG. Then, ten sub-models for deep ensemble learning take the inputs and predict the prognosis of LGG patients. While each prediction value is distinct among the sub-modules, the averaged values of these predictions can generalize the prediction and stabilize the random effects of deep learning.

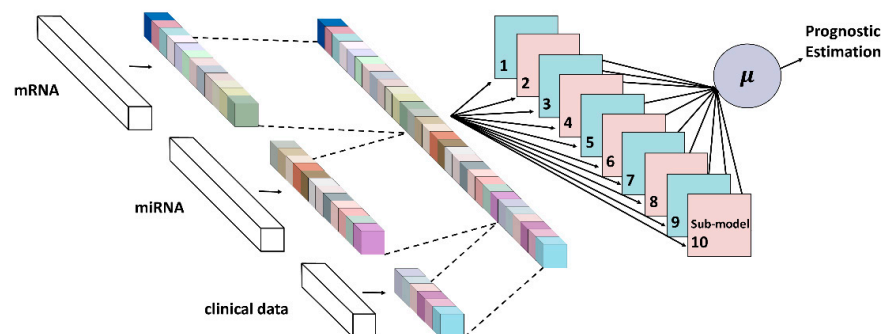


Figure 1. The deep learning architecture of Multi-PEN. The second layers with colors indicate the gene attention layers. The data multiplied by the attention values are transferred to sub-models. The estimation is the average value of each model.

Each sub-model consists of residual networks, as is shown in Figure 2A. Two residual blocks are used in the sub-models, each of which is composed of two fully connected layers, as is shown in Figure 2B. Layer normalization and the dropout technique are utilized between the layers to reduce the overfitting problem. The predictions of each sub-model are calculated with the hyperbolic tangent activation function, where the value ranges from -1.0 to 1.0 . A higher value indicates a high likelihood of survival, similar to GAENET (because the same loss function is used).

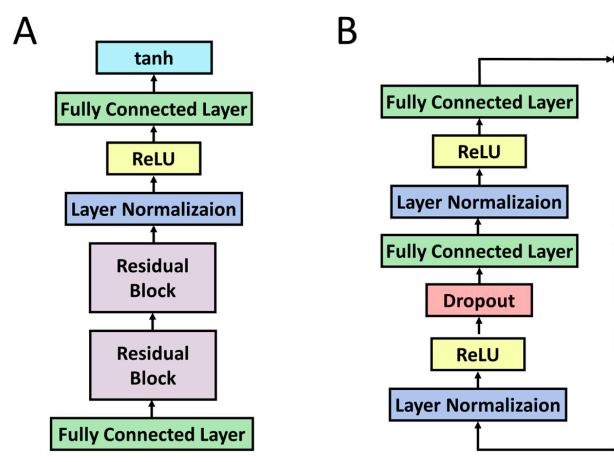


Figure 2. The detailed architecture of the sub-models in Multi-PEN. (A) The sub-model structures. (B) The structure of each residual block in the sub-models. The structure of the residual block in (A) is described in (B). The equivalent role organizing in (A,B) is colored identically.

3.2. Clinical Data

Multi-modal is used to refer to a framework that utilizes various datatypes and contexts. Multi-PEN employs a multi-modal framework with gene expression data and clinical data in order to enhance the estimation performance and to discover prognostic mRNA and miRNA while considering clinical data. Clinical variables used in the model are the age at diagnosis, sex, race, ethnicity, tumor grade, and the number of mutations. Isocitrate dehydrogenase (IDH) code subtypes are also included in the variables since it has been posited that IDH1 and 2 are significant prognostic and therapeutic biomarkers for glioma. Furthermore, 1p19q codeletion is also used for the same reason [48]. These clinical data are preprocessed to be used in Multi-PEN. The preprocessing methods are described in Section 3.1.

3.3. TCGA Data—Low-Grade Glioma

To evaluate Multi-PEN, The Cancer Genome Atlas for Low-Grade Glioma (TCGA-LGG) dataset was used in this study. In this experiment, 125 uncensored samples are employed to perform five-fold cross-validation; among the samples, 100 samples are randomly selected to train the model, and the other samples become test sets in each cross-validation. This five-fold cross-validation is performed 20 times; therefore, 100 experiments with different training and test sets are conducted [49]. For the preprocessing process, if the zero value of a gene exceeds half of the samples, the gene is omitted from the dataset. Additionally, if the samples have a missing value, the corresponding genes are excluded. The gene expression data are preprocessed by log normalization.

3.4. K-Fold Cross-Validation

K-fold cross-validation is used in various domains, including the medical domain, in order to make a dataset similar to the conditions in clinical tests [49–51]. K-fold cross-validation is a principled method for dividing a dataset into a training dataset and a validation dataset. Then, a machine learning model trains only with the training dataset, without any information in the validation set, including the number of samples and value distributions of each variable. In K-fold cross-validation, the dataset is equally and randomly divided into K sub-datasets to give the same number of samples. In this division process, the sub-datasets do not share any information, which gives identical conditions to clinical tests; therefore, the normalization process and preprocessing of the data are conducted only with the training set after the division; the validation set is preprocessed with the information of the training set. Then, $K - 1$ sub-datasets are used as the training set, and the other dataset is used as the validation set. This trial is repeated K times with different validation sets in each trial. Such processes of K-fold cross-validation enable a similar condition to K repetitions of clinical tests with different test sets.

3.5. Hyperparameter

In the experiment, Multi-PEN employs a total of ten sub-models. In the residual block, the number of nodes in the first and second fully connected layers is 16 and 8, respectively. Having such a small number of nodes in the fully connected layers is a result of the overfitting issue, where the number of samples is extremely small compared to the number of genes. The number of the attention layer nodes is set to 10 for each of the three attention layers. An optimization algorithm known as Adam is used, with the learning rate set to 0.0001. In this study, 20 epochs are used to train the model, and five mini-batches are organized in each epoch. For the performance comparison, the C-Index [52] is measured with the proposed model and other baseline models.

4. Results

A comparison of the performance measured by the C-Index of Multi-PEN, ridge regression, SurvivalSVM, RSF, and Coxnet is shown in Figure 3. With 20 times five-fold cross-validation, the C-Indices were evaluated 100 times. The C-Index of Multi-PEN was 0.7018.

While RSF exhibited effectiveness in terms of performance, Multi-PEN outperformed ridge and Coxnet, for which the performance was 0.6886 and 0.6771, respectively. In summary, Multi-PEN demonstrated competitive performance with other conventional methods.

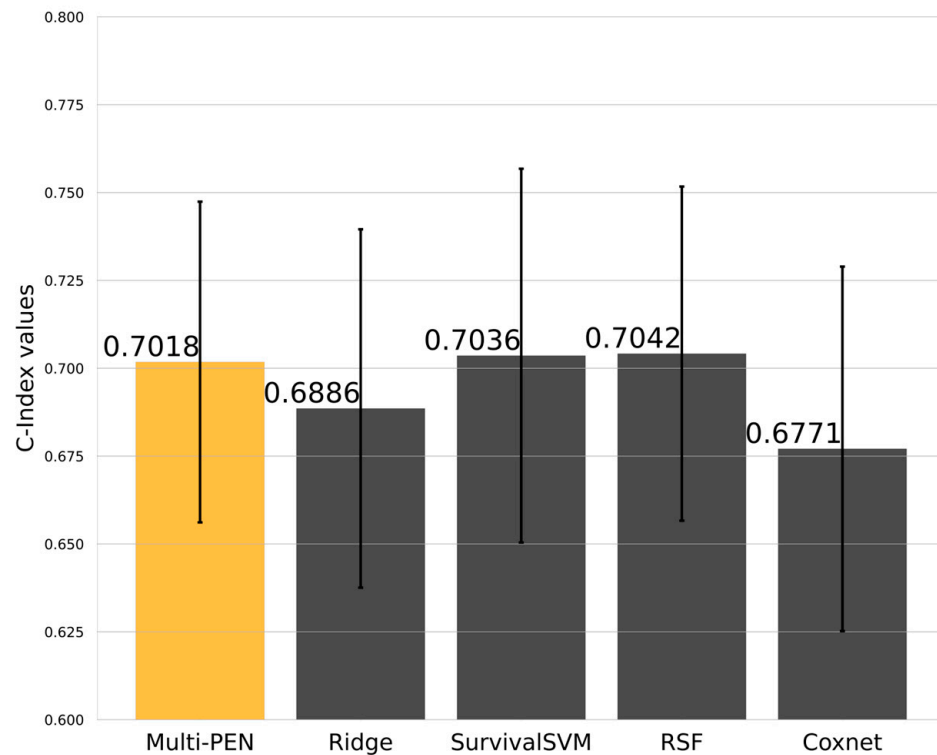


Figure 3. C-Index of Multi-PEN and other conventional models with 100 experiments.

As shown in Figure 3, the performance differences between the models were marginal, where the best model and the worst model exhibited a 4.0% difference in terms of the C-Index. Although Multi-PEN did not display outstanding performance compared to the conventional models, the performance difference between the best model and Multi-PEN was 0.3%, which can be interpreted as a competitive performance. Since deep learning models generally demonstrate better performances as the number of training samples increases, it is expected that this limitation can be handled in future studies with expanded samples.

The prognostic ranks of mRNA and miRNA can be found in the gene data using Multi-PEN, which is one of the main advantages of the model. This can lead us to the discovery of prognostic mRNA and miRNA. The values of the attention layer, which are shown in Figure 4, are where the prognostic genes of mRNA are located. Figure 5 also displays the fifty miRNA genes that may be used to predict outcomes. MYBL1 was shown to be the mRNA with the highest attention value, followed by CARHSP1, C9orf50, ABCC3, and HAX1. In addition, the prognostic miRNAs that were discovered by Multi-PEN are as follows: hsa-mir-421, hsa-mir-885, hsa-mir-495, hsa-mir-194-2, and hsa-mir-30d.

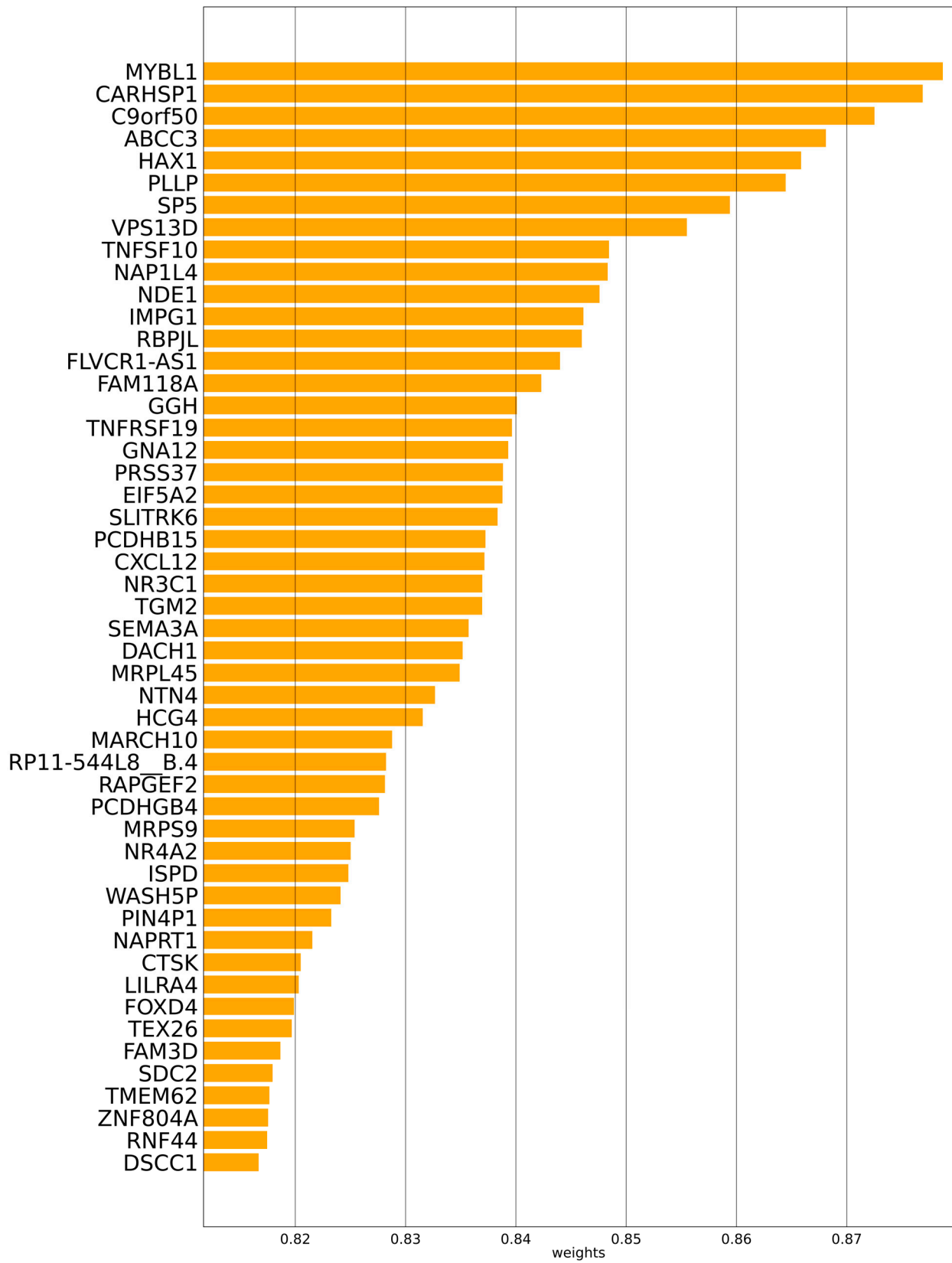


Figure 4. Estimated prognostic mRNAs by Multi-PEN and their attention values.

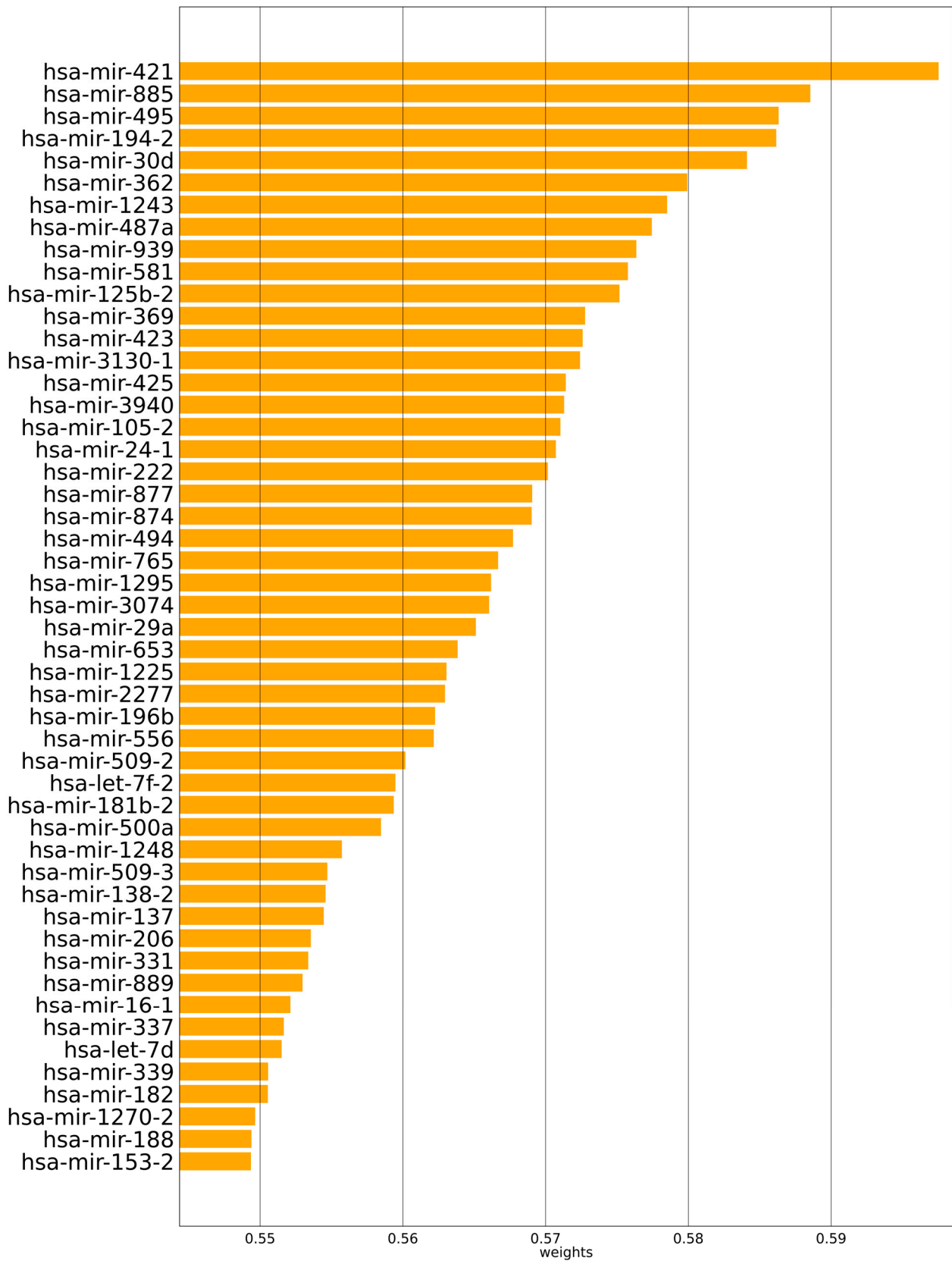


Figure 5. Estimated prognostic miRNAs by Multi-PEN and their attention values.

5. Discussion

The proposed deep learning model, Multi-PEN, was used to estimate prognosis and search for prognostic genes in LGG patients. Other studies have utilized conventional and straightforward deep learning models, such as MLP, without considering current developments in deep learning [21,22]. Conversely, to find the prognostic genes, including miRNA, Multi-PEN exploits recent developments, i.e., residual networks and gene attention mechanisms. The prognostic mRNA and miRNA were estimated using the gene attention mechanism. The prognosis prediction performance of Multi-PEN was competitive with other models, which suggests that the estimated prognostic mRNA and miRNA are reliable.

This study proposed a novel way of investigating prognostic mRNA and miRNA from a nonlinear perspective, while other studies have focused on linear perspectives with the Cox proportional-hazards model and t-test. The estimation of the prognostic genes can be conducted by a deep-learning-based attention algorithm embedded in Multi-PEN. Because prognostic genes are evaluated with combinational effects and polynomial correlations, gene attention can be considered to find genes from a nonlinear perspective. In contrast, the Cox proportional-hazards model and Kaplan–Meier estimation, which are commonly used in a variety of studies, are focused on the linear relationship between genes and outcomes.

MYBL1 was estimated to be the most significant mRNA with a gene attention value of approximately 0.88. This indicates that MYBL1 has a probability of possessing a direct relationship with the prognosis of LGG. Such an estimation of Multi-PEN is in accordance with several existing studies that have suggested that MYBL1 regulates cell survival, proliferation, and differentiation [53], which are highly associated with tumorigenesis.

Additionally, CARHSP1 was identified as a significant prognostic mRNA with the second-highest attention value. This result is also related to a recent study [54] that investigated whether CARHSP1 is radiotherapy protective in glioblastoma by signaling via the CARHSP1/TNF- α pathway. In addition, patients with high levels of CARHSP1 who are treated with radiotherapy have a poor prognosis. It was reported that ABCC3, one of the other prognostic mRNAs estimated by Multi-PEN, has a connection with a poor prognosis and resistance to treatment in cancer [55]. Additionally, according to the study, the intracellular concentration and efficacy of drugs are decreased by the exporter that releases the drugs from within the cells; thus, overexpressed ABCC3 creates resistance to multiple drugs for cancers, which may affect the prognosis of LGG. Another estimated prognostic mRNA, HAX1, has been reported to be related to lung cancers; HAX1 was overexpressed in non-small cell lung cancer; however, it was found that HAX1 was not overexpressed in normal cells [56]. Additionally, the high concentration of HAX1 gene expression was related to the TNM stage, lymphatic metastasis, and tumor size. While HAX1 has been previously reported as being related to lung cancers, the estimation of Multi-PEN implies that the gene also has a probability of being associated with LGG. Conversely, among the top five prognostic mRNAs estimated by Multi-PEN, C9orf50 has not been extensively studied regarding its relationship with LGG or other cancers. Further studies should be conducted on the relationship between C9orf50 and LGG, since the gene has a high attention value according to Multi-PEN.

Furthermore, prognostic miRNAs were also identified by Multi-PEN, which considered the relationship with mRNAs and other miRNAs, resulting from the gene attention layers in Multi-PEN. As a result, hsa-mir-421 was estimated to be the most significant miRNA with an attention value of 0.60; other estimated prognostic miRNAs were hsa-mir-885, hsa-mir-495, hsa-mir-194-2, and hsa-mir-30d. These results accord with those of existing studies demonstrating the overexpression of hsa-mir-421 in a variety of cancers, including gastric cancer [57,58], neuroblastoma [59], lung adenocarcinoma [60], breast cancer [61], and osteosarcoma [62]. Additionally, a study showed that has-mir-421 was highly expressed in lung cancer cells [63]. Furthermore, it has been demonstrated that the overexpression of has-mir-421 is a prognostic biomarker for non-small cell lung cancer and encourages tumor growth [64]. Considering these findings from existing studies, the estimation results by Multi-PEN are supported.

Additionally, another prognostic miRNA estimated by Multi-PEN, hsa-mir-885-5p, was reported to be upregulated in colorectal cancer cells [65]. Accordingly, the suppression of has-mir-885 restrains tumor cells from proliferating or migrating. Furthermore, it is known that abnormal expression of has-mir-885-5p has a possibility of reconstructing hepatocellular carcinoma metabolism and progression [66]. The direct relationship between hsa-mir-885 and LGG has not yet been investigated; however, it can be considered that hsa-mir-885 is associated with LGG based on these existing studies on hsa-mir-885 expression in cancer cells. Another estimated miRNA, hsa-mir-495, has been known as a tumor-suppressor [67]. Additionally, it was found that has-mir-495 is an oncogene [68]. Another miRNA, hsa-mir-194, has been linked to tumor metastasis and medication resistance in non-small cell lung cancer (NSCLC); hsa-mir-194 levels declined in NSCLC samples when compared to non-cancerous lung tissues, and a low hsa-mir-194 expression was associated with a poor prognosis [69]. For hsa-mir-30d, high expression contributed to invasion, proliferation, and tumor growth, resulting in poor survival of prostate cancer patients [70]. In addition, hsa-mir-30d overexpression inhibits the activation of cyclin D1, cMyc, and catenin, three essential components of the Wnt/beta-catenin signaling pathway, thereby inhibiting the development of colorectal cancer [71]. In contrast, in pancreatic cancer, hsa-mir-30d was significantly downregulated compared to non-tumor tissues and cells [72]. The capability of Multi-PEN to identify prognostic genes is supported by these existing studies. In addition, potential prognostic genes for LGG have been identified in this study and can be investigated further in future research.

6. Conclusions

Multi-PEN, a deep learning model with multi-omics and multi-modal schemes, was proposed in this paper to estimate the prognosis of LGG and prognostic genes. The general structure of Multi-PEN is similar to that of GAENET; however, the attention layers for miRNA expression and clinical data were additionally employed in Multi-PEN. Owing to the advantage of the gene attention mechanism, it can be used to identify prognostic mRNA and miRNA.

As a result, several mRNAs were estimated as prognostic mRNAs, including MYBL1, CARHSP1, C9orf50, ABCC3, and HAX1. Existing studies support these findings, as associations between these mRNAs and LGG and other cancers have been reported. In terms of prognostic miRNAs, hsa-mir-421, has-mir-885, has-mir-495, hsa-mir-194-2, and hsa-mir-30d were identified as the most significant. These miRNAs also accord with existing studies that have reported associations between these miRNAs and tumors. Using the proposed Multi-PEN, it is expected that the estimation of both prognosis and prognostic genes can be performed. Furthermore, it is anticipated that future studies can expand the use of Multi-PEN to other cancer types.

However, there was a limitation in this study in that the validation dataset was not composed of fresh samples with independent datasets; instead, K-fold cross-validation was employed for the evaluation. While it is expected that this experimental setting with a single dataset using K-fold cross-validation properly evaluates the performance of the models, the validation with independent datasets can elaborate the experiments in a more precise manner. Additionally, the randomness of the weight parameter initialization in deep learning models produces moderately different results for each experimental test. Although this limitation was partially addressed by 100 repetitions of experiments and examining average performance, other deep learning techniques may be required to handle this problem fundamentally.

Author Contributions: Conceptualization, S.R.C. and M.L.; methodology, S.R.C. and M.L.; software, S.R.C.; validation, S.R.C.; formal analysis, S.R.C.; investigation, S.R.C. and M.L.; writing—original draft preparation, S.R.C. and M.L.; supervision, M.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the Chung-Ang University Graduate Research Scholarship in 2022 as well as the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. NRF-2021R1F1A1050977).

Institutional Review Board Statement: Ethical review and approval were waived for this study, due to the nature of study and the application of a public dataset.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset employed in this study is a public dataset which is downloaded on 27 February 2022, received at Genomic Data Commons Data Portal: <https://portal.gdc.cancer.gov/> (accessed on 30 August 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Lee, M.; Tae, D.; Choi, J.H.; Jung, H.-Y.; Seok, J. Improved recurrent generative adversarial networks with regularization techniques and a controllable framework. *Inf. Sci.* **2020**, *538*, 428–443. [[CrossRef](#)]
2. Lee, M.; Seok, J. Estimation with uncertainty via conditional generative adversarial networks. *Sensors* **2021**, *21*, 6194. [[CrossRef](#)]
3. Park, M.; Lee, M.; Yu, S. HRGAN: A generative adversarial network producing higher-resolution images than training sets. *Sensors* **2022**, *22*, 1435. [[CrossRef](#)]
4. Alam, N.A.; Ahsan, M.; Based, A.; Haider, J.; Kowalski, M. COVID-19 detection from chest x-ray images using feature fusion and deep learning. *Sensors* **2021**, *21*, 1480. [[CrossRef](#)]
5. Yu, K.; Tan, L.; Lin, L.; Cheng, X.; Yi, Z.; Sato, T. Deep-learning-empowered breast cancer auxiliary diagnosis for 5GB remote E-health. *IEEE Wirel. Commun.* **2021**, *28*, 54–61. [[CrossRef](#)]
6. Liu, X.; Wang, H.; Li, Z.; Qin, L. Deep learning in ECG diagnosis: A review. *Knowl.-Based Syst.* **2021**, *227*, 107187. [[CrossRef](#)]
7. Liu, X.; Song, L.; Liu, S.; Zhang, Y. A review of deep-learning-based medical image segmentation methods. *Sustainability* **2021**, *13*, 1224. [[CrossRef](#)]
8. Li, Y.; Zhao, J.; Lv, Z.; Li, J. Medical image fusion method by deep learning. *Int. J. Cogn. Comput. Eng.* **2021**, *2*, 21–29. [[CrossRef](#)]
9. Kim, M.; Lee, B.-D. Automatic lung segmentation on chest X-rays using self-attention deep neural network. *Sensors* **2021**, *21*, 369. [[CrossRef](#)]
10. Xie, Y.; Zaccagna, F.; Rundo, L.; Testa, C.; Agati, R.; Lodi, R.; Manners, D.N.; Tonon, C. Convolutional neural network techniques for brain tumor classification (from 2015 to 2022): Review, challenges, and future perspectives. *Diagnostics* **2022**, *12*, 1850. [[CrossRef](#)]
11. Vitsios, D.; Dhindsa, R.S.; Middleton, L.; Gussow, A.B.; Petrovski, S. Prioritizing non-coding regions based on human genomic constraint and sequence context with deep learning. *Nat. Commun.* **2021**, *12*, 1–14. [[CrossRef](#)]
12. Liu, Y.; Huang, Y.; Wang, G.; Wang, Y. A deep learning approach for filtering structural variants in short read sequencing data. *Briefings Bioinform.* **2020**, *22*, bbaa370. [[CrossRef](#)]
13. Zhang, J.X.; Yordanov, B.; Gaunt, A.; Wang, M.X.; Dai, P.; Chen, Y.-J.; Zhang, K.; Fang, J.Z.; Dalchau, N.; Li, J.; et al. A deep learning model for predicting next-generation sequencing depth from DNA sequence. *Nat. Commun.* **2021**, *12*, 1–10. [[CrossRef](#)]
14. Vale-Silva, L.A.; Rohr, K. Long-term cancer survival prediction using multimodal deep learning. *Sci. Rep.* **2021**, *11*, 1–12. [[CrossRef](#)]
15. Allegra, A.; Tonacci, A.; Sciacotta, R.; Genovese, S.; Musolino, C.; Pioggia, G.; Gangemi, S. Machine Learning and Deep Learning Applications in Multiple Myeloma Diagnosis, Prognosis, and Treatment Selection. *Cancers* **2022**, *14*, 606. [[CrossRef](#)]
16. Arya, N.; Saha, S. Multi-modal advanced deep learning architectures for breast cancer survival prediction. *Knowledge-Based Syst.* **2021**, *221*, 106965. [[CrossRef](#)]
17. Hirtz, A.; Lebourdais, N.; Thomassin, M.; Rech, F.; Dumond, H.; Dubois-Pot-Schneider, H. Identification of Gender- and Subtype-Specific Gene Expression Associated with Patient Survival in Low-Grade and Anaplastic Glioma in Connection with Steroid Signaling. *Cancers* **2022**, *14*, 4114. [[CrossRef](#)]
18. Louis, D.N.; Ohgaki, H.; Wiestler, O.D.; Cavenee, W.K.; Burger, P.C.; Jouvet, A.; Scheithauer, B.W.; Kleihues, P. The 2007 WHO classification of tumours of the central nervous system. *Acta Neuropathol.* **2007**, *114*, 97–109. [[CrossRef](#)]
19. Forst, D.A.; Nahed, B.V.; Loeffler, J.S.; Batchelor, T.T. Low-grade gliomas. *Oncologist* **2014**, *19*, 403–413.
20. Hu, Y.; Deng, F.; Zhang, L.; Hu, K.; Liu, S.; Zhong, S.; Yang, J.; Zeng, X.; Peng, X. Depression and Quality of Life in Patients with Gliomas: A Narrative Review. *J. Clin. Med.* **2022**, *11*, 4811. [[CrossRef](#)]
21. Wang, D.; Ge, L.; Guo, Z.; Li, Y.; Zhu, B.; Wang, W.; Wei, C.; Li, Q.; Wang, Z. Efficacy and safety of trametinib in neurofibromatosis type 1-associated plexiform neurofibroma and low-grade glioma: A systematic review and meta-analysis. *Pharmaceuticals* **2022**, *15*, 956. [[CrossRef](#)]
22. Pasquini, L.; Jenabi, M.; Yildirim, O.; Silveira, P.; Peck, K.K.; Holodny, A.I. Brain functional connectivity in low- and high-grade gliomas: Differences in network dynamics associated with tumor grade and location. *Cancers* **2022**, *14*, 3327. [[CrossRef](#)]
23. Lombard, A.; Duffau, H. Sexual dysfunction of patients with diffuse low-grade glioma: A qualitative review of a neglected concern. *Cancers* **2022**, *14*, 3025. [[CrossRef](#)]

24. Claus, E.B.; Walsh, K.M.; Wiencke, J.K.; Molinaro, A.M.; Wiemels, J.L.; Schildkraut, J.M.; Bondy, M.L.; Berger, M.; Jenkins, R.; Wrensch, M. Survival and low-grade glioma: The emergence of genetic information. *Neurosurg. Focus* **2015**, *38*, E6. [[CrossRef](#)]
25. Bettgowda, C.; Agrawal, N.; Jiao, Y.; Sausen, M.; Wood, L.D.; Hruban, R.H.; Rodriguez, F.J.; Cahill, D.P.; McLendon, R.; Riggins, G.; et al. Mutations in *CIC* and *FUBP1* contribute to human oligodendroglioma. *Science* **2011**, *333*, 1453–1455. [[CrossRef](#)]
26. Sahm, F.; Koelsche, C.; Meyer, J.; Pusch, S.; Lindenberg, K.; Mueller, W.; Herold-Mende, C.; Von Deimling, A.; Hartmann, C. *CIC* and *FUBP1* mutations in oligodendrogliomas, oligoastrocytomas and astrocytomas. *Acta Neuropathol.* **2012**, *123*, 853–860. [[CrossRef](#)]
27. Lee, M. An ensemble deep learning model with a gene attention mechanism for estimating the prognosis of low-grade glioma. *Biology* **2022**, *11*, 586. [[CrossRef](#)]
28. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 6000–6010.
29. Tay, Y.; Bahri, D.; Metzler, D.; Juan, D.-C.; Zhao, Z.; Zheng, C. Synthesizer: Rethinking self-attention for transformer models. In Proceedings of the 38th International Conference on Machine Learning (PMLR), Virtual, 18–24 July 2021; pp. 10183–10192.
30. Bazi, Y.; Bashmal, L.; Rahhal, M.M.A.; Dayil, R.A.; Ajlan, N.A. Vision transformers for remote sensing image classification. *Remote Sens.* **2021**, *13*, 516. [[CrossRef](#)]
31. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
32. Sagi, O.; Rokach, L. Ensemble learning: A survey. *Wiley Interdiscip. Rev.* **2018**, *8*, e1249. [[CrossRef](#)]
33. Wu, W.; Wang, Y.; Xiang, J.; Li, X.; Wahafu, A.; Yu, X.; Bai, X.; Yan, G.; Wang, C.; Wang, N.; et al. A novel multi-omics analysis model for diagnosis and survival prediction of lower-grade glioma patients. *Front. Oncol.* **2022**, *12*, 729002. [[CrossRef](#)]
34. Kazerooni, A.F.; Saxena, S.; Toorens, E.; Tu, D.; Bashyam, V.; Akbari, H.; Mamourian, E.; Sako, C.; Koumenis, C.; Verginadis, I.; et al. Clinical measures, radiomics, and genomics offer synergistic value in AI-based prediction of overall survival in patients with glioblastoma. *Sci. Rep.* **2022**, *12*, 1–13. [[CrossRef](#)]
35. Pan, X.; Burgman, B.; Wu, E.; Huang, J.H.; Sahni, N.; Yi, S.S. i-Modern: Integrated multi-omics network model identifies potential therapeutic targets in glioma by deep learning with interpretability. *Comput. Struct. Biotechnol. J.* **2022**, *20*, 3511–3521. [[CrossRef](#)]
36. O'Brien, J.; Hayder, H.; Zayed, Y.; Peng, C. Overview of microRNA biogenesis, mechanisms of actions, and circulation. *Front. Endocrinol.* **2018**, *9*, 402. [[CrossRef](#)]
37. Li, Y.; Wang, J.; Elzo, M.; Gan, M.; Tang, T.; Shao, J.; Lai, T.; Ma, Y.; Jia, X.; Lai, S. Multi-Omics Analysis of Key microRNA–mRNA Metabolic Regulatory Networks in Skeletal Muscle of Obese Rabbits. *Int. J. Mol. Sci.* **2021**, *22*, 4204. [[CrossRef](#)]
38. Kim, Y.; Lee, S.; Jang, J.-Y.; Lee, S.; Park, T. Identifying miRNA–mRNA Integration Set Associated with Survival Time. *Frontiers in Genetics* **2021**, *12*, 634922. [[CrossRef](#)]
39. Amal, S.; Safarnejad, L.; Omiye, J.A.; Ghanzouri, I.; Cabot, J.H.; Ross, E.G. Use of Multi-Modal Data and Machine Learning to Improve Cardiovascular Disease Care. *Front. Cardiovasc. Med.* **2022**, *9*, 840262. [[CrossRef](#)]
40. Shrestha, A.; Mahmood, A. Review of deep learning algorithms and architectures. *IEEE Access* **2019**, *7*, 53040–53065. [[CrossRef](#)]
41. Zhou, T.; Lu, H.; Yang, Z.; Qiu, S.; Huo, B.; Dong, Y. The ensemble deep learning model for novel COVID-19 on CT images. *Appl. Soft Comput.* **2021**, *98*, 106885. [[CrossRef](#)]
42. Ba, J.L.; Kiros, J.R.; Hinton, G.E. Layer normalization. *ArXiv* **2016**, arXiv:06450. Preprint.
43. Pölstler, S.; Navab, N.; Katouzian, A. Fast training of support vector machines for survival analysis. In Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Porto, Portugal, 7–11 September 2015; pp. 243–259.
44. Ishwaran, H.; Kogalur, U.B.; Blackstone, E.H.; Lauer, M.S. Random survival forests. *Ann. Appl. Stat.* **2008**, *2*, 841–860. [[CrossRef](#)]
45. Simon, N.; Friedman, J.H.; Hastie, T.; Tibshirani, R. Regularization paths for cox's proportional hazards model via coordinate descent. *J. Stat. Softw.* **2011**, *39*, 1–13. [[CrossRef](#)]
46. Chaudhary, K.; Poirion, O.B.; Lu, L.; Garmire, L.X. Deep learning-based multi-omics integration robustly predicts survival in liver cancer using deep learning to predict liver cancer prognosis. *Clin. Cancer Res.* **2018**, *24*, 1248–1259. [[CrossRef](#)]
47. Huang, Z.; Zhan, X.; Xiang, S.; Johnson, T.S.; Helm, B.; Yu, C.Y.; Zhang, J.; Salama, P.; Rizkalla, M.; Han, Z.; et al. SALMON: Survival analysis learning with multi-omics neural networks on breast cancer. *Front. Genet.* **2019**, *10*, 166. [[CrossRef](#)]
48. Liu, S.; Shah, Z.; Sav, A.; Russo, C.; Berkovsky, S.; Qian, Y.; Coiera, E.; Di Ieva, A. Isocitrate dehydrogenase (IDH) status prediction in histopathology images of gliomas using deep learning. *Sci. Rep.* **2020**, *10*, 7711–7733. [[CrossRef](#)]
49. Prusty, S.; Patnaik, S.; Dash, S.K. SKCV: Stratified K-fold cross-validation on ML classifiers for predicting cervical cancer. *Front. Nanotechnol.* **2022**, *4*, 972421. [[CrossRef](#)]
50. Rafafo, M. Cross validation methods: Analysis based on diagnostics of thyroid cancer metastasis. *ICT Express* **2021**, *8*, 183–188. [[CrossRef](#)]
51. Normawati, D.; Ismi, D.P. K-fold cross validation for selection of cardiovascular disease diagnosis features by applying rule-based datamining. *Signal Image Process. Lett.* **2019**, *1*, 23–35. [[CrossRef](#)]
52. Lee, S.; Lim, H. Review of statistical methods for survival analysis using genomic data. *Genom. Inform.* **2019**, *17*, e41. [[CrossRef](#)]
53. Cicerò, Y.; Sala, A. MYB oncoproteins: Emerging players and potential therapeutic targets in human cancer. *Oncogenesis* **2021**, *10*, 1–15. [[CrossRef](#)]

54. Zhu, G.-D.; Yu, J.; Sun, Z.-Y.; Chen, Y.; Zheng, H.-M.; Lin, M.-L.; Ou-Yang, S.; Liu, G.-L.; Zhang, J.-W.; Shao, F.-M. Genome-wide CRISPR/Cas9 screening identifies CARHSP1 responsible for radiation resistance in glioblastoma. *Cell Death Dis.* **2021**, *12*, 1–9. [[CrossRef](#)]
55. Ramírez-Cosmes, A.; Reyes-Jiménez, E.; Zertuche-Martínez, C.; Hernández-Hernández, C.A.; García-Román, R.; Romero-Díaz, R.I.; Manuel-Martínez, A.E.; Elizarrarás-Rivas, J.; Vásquez-Garzón, V.R. The implications of ABCC3 in cancer drug resistance: Can we use it as a therapeutic target? *Am. J. Cancer Res.* **2021**, *11*, 4127.
56. Liang, Z.; Zhong, Y.; Meng, L.; Chen, Y.; Liu, Y.; Wu, A.; Li, X.; Wang, M. HAX1 enhances the survival and metastasis of non-small cell lung cancer through the AKT/mTOR and MDM2/p53 signaling pathway. *Thorac. Cancer* **2020**, *11*, 3155–3167. [[CrossRef](#)]
57. Wu, J.; Li, G.; Yao, Y.; Wang, Z.; Sun, W.; Wang, J. MicroRNA-421 is a new potential diagnosis biomarker with higher sensitivity and specificity than carcinoembryonic antigen and cancer antigen 125 in gastric cancer. *Biomarkers* **2014**, *20*, 58–63. [[CrossRef](#)]
58. Liu, H.; Gao, Y.; Song, D.; Liu, T.; Feng, Y. Correlation between microRNA-421 expression level and prognosis of gastric cancer. *Int. J. Clin. Exp. Pathol.* **2015**, *8*, 15128–15132.
59. Li, Y.; Li, W.; Zhang, J.-G.; Li, H.-Y.; Li, Y.-M. Downregulation of tumor suppressor menin by miR-421 promotes proliferation and migration of neuroblastoma. *Tumor Biol.* **2014**, *35*, 10011–10017. [[CrossRef](#)]
60. Cinegaglia, N.C.; Andrade, S.C.S.; Tokar, T.; Pinheiro, M.; Severino, F.E.; Oliveira, R.A.; Hasimoto, E.N.; Cataneo, D.C.; Cataneo, A.J.; Defaveri, J. Integrative transcriptome analysis identifies deregulated microRNA-transcription factor networks in lung adenocarcinoma. *Oncotarget* **2016**, *7*, 28920–28934. [[CrossRef](#)]
61. Lou, W.; Liu, J.; Ding, B.; Jin, L.; Xu, L.; Li, X.; Chen, J.; Fan, W. Five miRNAs-mediated PIEZO2 downregulation, accompanied with activation of Hedgehog signaling pathway, predicts poor prognosis of breast cancer. *Aging* **2019**, *11*, 2628–2652. [[CrossRef](#)]
62. Zhou, S.; Wang, B.; Hu, J.; Zhou, Y.; Jiang, M.; Wu, M.; Qin, L.; Yang, X. miR-421 is a diagnostic and prognostic marker in patients with osteosarcoma. *Tumor Biol.* **2016**, *37*, 9001–9007. [[CrossRef](#)]
63. Mo, X.-M.; Qin, P.-F.; Wang, B.; Liu, F.-H.; Li, H.-H. miR-421 promotes the viability of A549 lung cancer cells by targeting forkhead box O1. *Oncol. Lett.* **2020**, *20*, 306. [[CrossRef](#)]
64. Dong, A.; Zhang, J.; Sun, W.; Hua, H.; Sun, Y. Upregulation of miR-421 predicts poor prognosis and promotes proliferation, migration, and invasion of papillary thyroid cancer cells. *J. Chin. Med. Assoc.* **2020**, *83*, 991–996. [[CrossRef](#)]
65. Su, M.; Qin, B.; Liu, F.; Chen, Y.; Zhang, R. miR-885-5p upregulation promotes colorectal cancer cell proliferation and migration by targeting suppressor of cytokine signaling. *Oncol. Lett.* **2018**, *16*, 65–72. [[CrossRef](#)]
66. Xu, F.; Yan, J.-J.; Gan, Y.; Chang, Y.; Wang, H.-L.; He, X.-X.; Zhao, Q. miR-885-5p Negatively Regulates Warburg Effect by Silencing Hexokinase 2 in Liver Cancer. *Mol. Ther. Nucleic Acids* **2019**, *18*, 308–319. [[CrossRef](#)]
67. Jiang, X.; Huang, H.; Li, Z.; He, C.; Li, Y.; Chen, P.; Gurbuxani, S.; Arnovitz, S.; Hong, G.-M.; Price, C.; et al. miR-495 is a tumor-suppressor microRNA down-regulated in MLL-rearranged leukemia. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 19397–19402. [[CrossRef](#)]
68. Chen, H.; Wang, X.; Bai, J.; He, A. Expression, regulation and function of miR-495 in healthy and tumor tissues. *Oncol. Lett.* **2017**, *13*, 2021–2026. [[CrossRef](#)]
69. Zhu, X.; Li, D.; Yu, F.; Jia, C.; Xie, J.; Ma, Y.; Fan, S.; Cai, H.; Luo, Q.; Lv, Z.; et al. miR-194 inhibits the proliferation, invasion, migration, and enhances the chemosensitivity of non-small cell lung cancer cells by targeting forkhead box A1 protein. *Oncotarget* **2016**, *7*, 13139–13152. [[CrossRef](#)]
70. Lin, Z.Y.; Chen, G.; Zhang, Y.Q.; He, H.C.; Liang, Y.X.; Ye, J.H.; Liang, Y.K.; Mo, R.J.; Lu, J.M.; Zhuo, Y.J. MicroRNA-30d promotes angiogenesis and tumor growth via MYPT1/c-JUN/VEGFA pathway and predicts aggressive outcome in prostate cancer. *Mol. Cancer* **2017**, *16*, 1–14. [[CrossRef](#)]
71. Yan, L.; Qiu, J.; Yao, J. Downregulation of microRNA-30d promotes cell proliferation and invasion by targeting LRR-1 in colorectal carcinoma. *Int. J. Mol. Med.* **2017**, *39*, 1371–1380. [[CrossRef](#)]
72. Xu, X.; Zong, K.; Wang, X.; Dou, D.; Lv, P.; Zhang, Z.; Li, H. miR-30d suppresses proliferation and invasiveness of pancreatic cancer by targeting the SOX4/PI3K-AKT axis and predicts poor outcome. *Cell Death Dis.* **2021**, *12*, 1–14. [[CrossRef](#)]