Multimodal data integration for ovarian cancer overall survival prediction

Bucciarelli Davide, Pavone Cosimo, Salici Giacomo
{271102,272636,270385}@studenti.unimore.it

February 2024

Abstract

In the field of oncology, Overall Survival (OS, time between the diagnosis and the death of a patient) is a crucial piece of information used to schedule follow-up medical examinations and to define the correct type of treatment according to the characteristics and the specific conditions of each patient. (personalized treatment). The prediction of overall survival depends simultaneously on information coming both from the genetic characteristics of the tumor and from the appearance of the cancer-affected tissue. To date, most overall survival prediction models mainly use molecular data (multi-omics). Still, simultaneously, integrating molecular data and histological images could be advantageous. Our work analyzes the SOTA models in both tasks and then creates a novel integration method to maximize the interaction between the different information contained in the modalities. In particular, two architectures have been explored: a Siamese network that tries to separate specific and shared information, and an autoencoder that is aimed to minimize the hazard risk error. We were able to reach an accuracy of 0.69 in overall survival binary classification. Code is publicly available on github.com/cosimop2000/Multimodal-data-integration-for-OVCA-survival-prediction

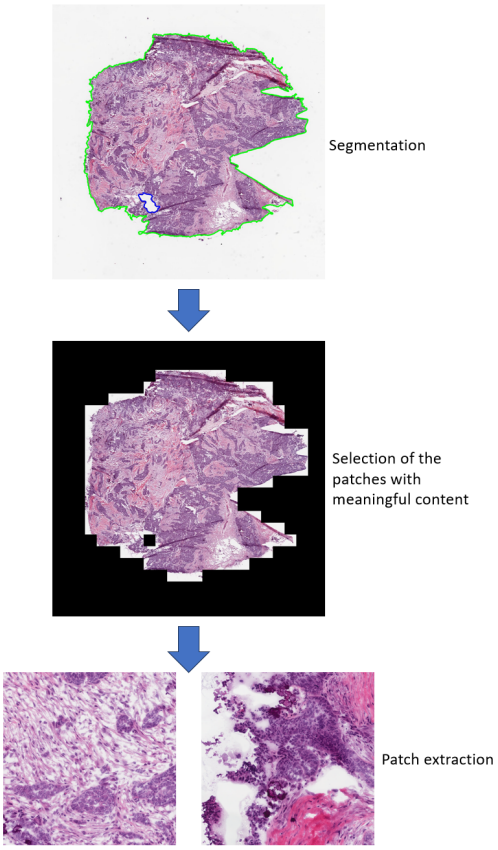


Figure 1: Whole Slide Images preprocessing

1 Introduction and related work

Predicting the survival of cancer patients has posed an intriguing and complex challenge in cancer research for the last few decades [1]. The heterogeneous nature of cancer, marked by diverse molecular features, clinical behaviors, morphological appearances, and responses to therapies, adds complexity to the task. Thus, the integration of multiple data sources is fundamental to have enough information to describe the heterogeneity of ovarian cancer disease [2]. A very common challenge that every paper must address is the so-called '*curse of dimensionality*' that leads the models to drastic overfitting when a highly dimensional dataset is given as an input [3][4]. This happens because ovarian cancer data is often poorly available because of the scarcity of the phenotype. At the same time, we have to deal with tens of thousands of genomic and histological features, that can be extremely noisy and misleading.
To address this problem we exploited data fusion techniques already available in literature [5][6], which we slightly modified according to our needs, to include both multi-omics data and WSI images. In particular, we dealt with WSI images by first pre-processing and obtaining patches with CLAM [7], then we extracted relevant features in a self-supervised way using PLIP [8] (Figure 1). Then, we introduced a random patch selection technique to try to mitigate the effects of overfitting, as well as a zero-shot classification with the already mentioned model that aims to select the most interesting patches. For the multi-omics part, we used DNA methylation values, gene expression data (mRNA), and copy number variation (CNV).

In the literature, several methods are proposed to

	BRCA	BLCA
SurvPath	0.655	0.625
MCAT	0.580	0.624
Customics	0.633	0.637

Table 1: Comparison between the C-index values between several models found in literature and two cohorts.

solve the multi-omics integration. According to this systematic review [9], WSIs are not always used, and in our experiments we conducted ablative studies to quantify their impact. Regarding the base models, both deep neural networks and classical machine learning-based networks are used. The latter ones are less frequent nowadays but are still employed, in particular, [10] gene prediction is made using support vector machines (SVM) and logistic regression (LR). Other, more complex, multi-modal fusion models are based on CNN,LSTM and Transformers [11] [12]. Finally, in the last year, models based on autoencoders [5] [13] [14] or based on custom networks, such as siamese network [6], have been developed. These last ones are the works more closely related to our task, so we decided to start from these frameworks and adapt them to our needs.

The main metric used for evaluating the OS prediction performance is the **C-Index** [15], as shown in table 1. It serves as a better extension of the area under the ROC curve (AUC). However, unlike AUC, the C-Index takes into account censored data, rendering it particularly suitable for survival analysis. It evaluates the model's discrimination power, that is the ability to provide a reliable ranking of the survival times based on the individual risk scores.

$$\text{C-index} = \frac{\sum_{i,j} \mathbb{1}_{T_j<T_i} \cdot \mathbb{1}_{\eta_j>\eta_i} \cdot \delta_j}{\sum_{i,j} \mathbb{1}_{T_j<T_i} \cdot \delta_j} \tag{1}$$

where η_i is the risk score of sample i, $\mathbb{1}_{T_j<T_i} = 1$ else 0, $\mathbb{1}_{\eta_j>\eta_i} = 1$ else 0. As with the AUC, a C-index equal to 1 means the best models, while a C-index of 0.5 is a random prediction.

2 Dataset

We collected data from the National Cancer Institute GDC data portal[1]. In particular, we selected the cases where the cancer's primary site is the ovary and coming from the *TCGA-OV* project, since they were the most numerous. In total, we had to deal with 600 cases. Of those, 349 were recorded when the death of the patient had unfortunately already happened, while 236 were still alive, and for 2 cases the vital status was unknown. To apply our methods, we had to use just the cases where all the modalities were available ensuring comprehensive integration and analysis of multimodal data for ovarian cancer research. In particular:

[1] https://portal.gdc.cancer.gov/

Methylation Beta values: In our study, we utilized data coming from the Illumina Human Methylation 27 platform with the SeSAMe Methylation Beta Estimation workflow to measure methylation levels at CpG sites as beta values.

mRNA: The mRNA gene expression data employed in this study were obtained through RNA-Seq technology, with read counts measured on a gene level utilizing the STAR (Spliced Transcripts Alignment to a Reference) aligner. To ensure accurate comparisons across samples, the original read counts were normalized using two commonly available methods: Fragments Per Kilobase of transcript per Million mapped reads (FPKM) and FPKM Upper Quartile (FPKM-UQ). FPKM-UQ normalization is a modified FPKM calculation that replaces the sequencing quantity with the protein-coding gene in the 75th percentile position. By using the upper quartile value instead of the total sequencing quantity, FPKM-UQ aims to mitigate the influence of noisy genes and provide a more stable measure of gene expression. Gene expression quantification focused exclusively on protein-coding genes.

CNV: Gene-level copy number data employed in this study were obtained from ASCAT3 workflows. Copy number variation values were derived as the weighted median of copy number values from all overlapped segments.

WSI: Whole slide histological images of ovarian cancer, encompassing cystic, mucinous, and serous neoplasms, offer a comprehensive visual depiction of the disease. The WSI utilized in this study were approximately 30000 x 30000 pixels in size. It's important to note that while these images provide considerable resolution, they may not match the highest quality standards observed in other studies. Nonetheless, they yielded valuable morphological (pertaining to physical characteristics and structures within the tissue sample, such as cell size, shape, and organization) and spatial information crucial for our analysis.

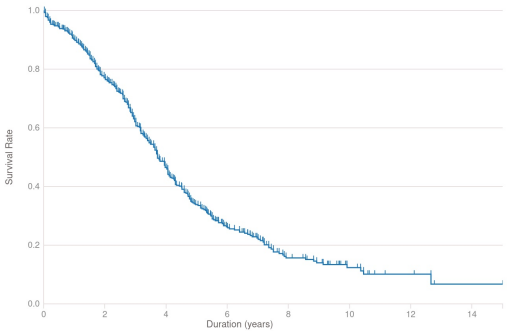


Figure 2: Overall Survival curve of our case selection

3 Methods

3.1 Preprocessing

The pipeline that we developed starts from a simple preprocessing of the omics data:

- Remove the features for which at least one sample had a null value
- Apply a standard scaler

As for the WSI images, the main problem was their great resolution, thus we had to perform a heavy processing step so that, at train time, we just had to deal with relatively small embeddings:

- Create 1024x1024 pixels non overlapping patches from the original WSI image, using CLAM to identify the relevant patches focusing on segmenting the tissue and excluding any holes. Each slide was processed to extract arrays of tissue patches, stored in separate `.h5` files, where each entry corresponds to the coordinates of the top-left corner of a patch. We employed a preset specifically designed for segmenting TCGA slides to adjust the process.
  - Segmentation: convert the WSI to the HSV color space. A binary mask for tissue regions is computed by thresholding the saturation channel after median blurring. Additional morphological operations fill gaps and holes. Approximate contours of foreground objects are filtered based on area and stored. Segmentation masks are provided for optional visual inspection.
  - Patching: generate 1024×1024 patches from within the segmented foreground contours, ensuring that the border of the slide is padded and verifying if all four points in a small grid around the center of the patch are inside the contour before cropping. These patches, along with their coordinates and slide metadata, are stored as stacks of image patches using the HDF5 hierarchical data format.
- For each patch, we extracted a feature representation using PLIP (Pathology Language and Image Pre-Training) [8], a large-scale pre-trained model that can be used to extract visual and language features from pathology images and text description. The model is a fine-tuned version of the original CLIP model. PLIP was trained on OpenPath, a large dataset of 208,414 pathology images paired with natural language descriptions. The image encoder leverages a Vision Transformer (ViT) to obtain a 512-dimensional feature vector for each input image.
- After creating the feature representations, we saved all the embeddings for all the images related to the patients. This allowed us to perform patch selection efficiently at a later stage of the analysis.

In particular, we tried two alternative methods for performing the selection of the most relevant patches to feed into our network:

Most relevant: We used PLIP to perform zero-shot classification on the patches, classifying them as either tumoral or not, then we selected the 10 patches with the highest scores for the 'tumor' class.

Random: Dynamic random selection of 10 patches. This way, each time an image is forwarded in the model, the patch mix is different, allowing the model to 'see' gradually most of the image during the training.

After this preprocessing step, we ended up with roughly:

- 19k features for the methylation
- 19k features for gene expression
- 5k features for the image feature vector

We then applied two methods to perform the overall survival prediction.

3.2 The autoencoder model

The first architecture that we considered to use as a backbone was a variational autoencoder (VAE). As we said in section 1, several methods in the literature consider this type of generative model. The great advantage is that the encoders can learn a meaningful latent space starting from high dimensional input data. In particular, the VAE encodes a distribution, typically Gaussian, since during the training the sampling is then regularized to avoid "unbalanced" latent space. Consequently, the loss function will have a double term, the first one ($\mathcal{L}_{\text{rec}}$) will focus on the ability to reconstruct the data, while the second one ($\mathcal{L}_{\text{kl}}$), typically a Kullback-Leibler divergence, is used to bring the encoded distribution as close as possible to the theoretical one.

$$\mathcal{L}_{\text{VAE}} = \mathcal{L}_{\text{rec}} + \mathcal{L}_{\text{kl}} = \\ - \mathbb{E}_{q_\phi(\mathbf{z}|\mathbf{x})} \left[\log p_\theta(\mathbf{x}|\mathbf{z})\right] - \mathbb{E}_{q_\phi} \left[- \log \frac{p_\theta(\mathbf{z})}{p_\theta(\mathbf{z}|\mathbf{x})} \right] \quad (2)$$

During our experiments, several VAE-based methods have been applied. In particular, firstly we studied both [5] and [14] that considered the problem of multi-omics integration. Both these models - instead of the simple Kullback-Leibler divergence-based loss - use a Maximum Mean Discrepancy loss (MMD, first proposed in [16]), that tries to solve the representation limit and the over-fitting risk of the first one. The MMD loss can be written as

$$\text{MMD}(p(z)||q(z)) = \mathbb{E}_{p(z),p(z')}[\mathcal{K}(z,z\prime)] + \mathbb{E}_{q(z),q(z')}[\mathcal{K}(z,z')] - 2\mathbb{E}_{p(z),q(z')}[\mathcal{K}(z,z')] \quad (3)$$

where $\mathcal{K}$ is the Gaussian kernel. MMD loss will substitute the $\mathcal{L}_{\text{kl}}$ term in (2).

To overcome the problem of joining different sources, [5] implemented a custom integration that makes use of a two-step architecture, as shown in figure 3. Differently from the well-known literature approaches such as joint integration (JI; where the representations are created inside the model) and late integration (LI; where each source is learned separately), the proposed mixed integration consists of learning first the sources independently - as in LI - and then joining them and finish the training considering them as a single input as in the JI.

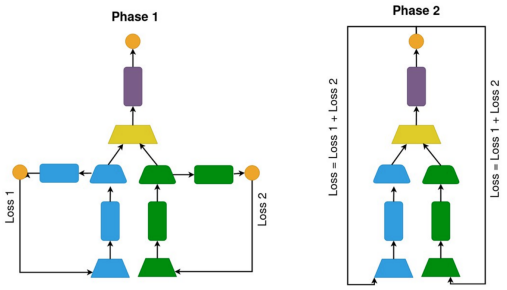


Figure 3: Mixed integration from [5]

Finally, the total loss is formed by the above-presented loss summed with a task-specific loss that in the case of the survival task may be taken from the DeepSurv hazard model [17]. In that model, the negative partial log-likelihood estimates the Cox risk function. To sum up,

$$\mathcal{L}_{\text{tot}} = \mathcal{L}_{\text{AE}} + \alpha\mathcal{L}_{\text{task}} \quad (4)$$

$$\mathcal{L}_{\text{task}}(\theta) = -\sum_{i:E_i=1}\left(\hat{\mu}(\mathbf{x}_i;\theta) - \log\sum_{j\in R(T_i)} e^{\hat{\mu}(\mathbf{x}_j;\theta)}\right) \quad (5)$$

The above-presented architecture worked well with our omics data but at the same time, we chose to continue using the following second method that is nearer to the current state-of-the-art, also because it is more feasible for an actual multi-modal integration with the WSI.

3.3 The Siamese model

This method consists of a slight alteration, shown in figure 4, of what was proposed in [6], where we process multi-omics data by trying to separate the pieces of information that are shared across all the omics and the information that is instead specific for each modality. This is done by training two separate networks for each input:

- A shared network, forced to learn similar representations across the various input modes thanks to the "distance" loss applied for the training of the shared net of all the inputs.
- A specific network, trained to learn a representation of the input that contains specific information not provided by the other modalities, thanks to an Orthogonal loss applied between the specific and the shared network of a single input.

In particular, the orthogonality constraint is defined as a loss function applied at every layer of the feature extraction:

$$\mathcal{L}_{\text{diff}} = ||H_k^c(m)^T H_k^s(m)||^2$$

Where $H_k^c(m)$ is the output of the m-th shared layer of the k-th modality, and $H_k^s(m)$ is the output of the m-th layer of the k-th modality. On the other hand, the distance constraint, applied at each layer between the shared representation of the different omics, is implemented as such:

$$d_n = ||h_{k,n}^c - h_{l,n}^c||^2$$
$$\mathcal{L}_{\text{diss}} = \frac{1}{2N}\sum_{n=1}^{N} d_n^2$$

Where $h_{k,n}^c$ and $h_{l,n}^c$ are the shared representations of sample x_n of modalities k, l.

The task performed in this case is binary classification of short and long time survivors, so, according to the distribution of the survival times, we selected a threshold of four years to distinguish between classes 0 and 1. Moreover, to be able to perform this kind of classification, we had to use just the cases for which the overall survival times were known. This means that we could only use the cases where the death of the patient had already occurred.

Then, the final feature vector is created by concatenating the outputs of all the specific networks and the shared ones. We then use this as the input to an MLP classifier. In the original paper, the classifier was trained using both a cross-entropy loss and a contrastive loss, but in our case, we decided to only use the cross-entropy since the optimization would end up being stuck with low classification results if we also included the contrastive objective.

The main novelty we introduced is to use the embedding of a WSI image as an input to the Siamese model. To create the embedding of the image, as we said before we came up with two different embedding selection methods. The first consists of blindly relying on a supposedly strong zero shot classifier to find the patches with the highest score for the class 'tumor', then selecting the 10 most relevant ones and concatenating their PLIP embedding to create the final WSI feature vector. The second method relies on the dynamic creation of the embedding of the images at train time, starting from the list of the patch embeddings created previously.

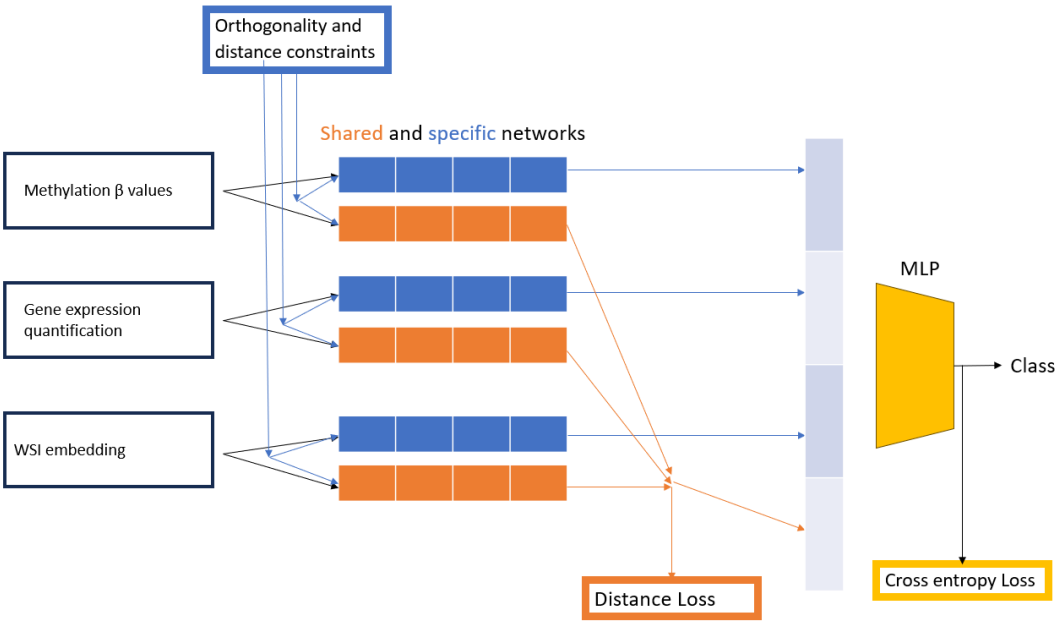


Figure 4: Scheme of the Siamese architecture

- We know that, for each patient, there could be multiple WSI images related to them, so we randomly select one.
- Then, we sample the embeddings of 10 patches coming from the selected image and concatenate them together.

This random selection method is done in order to try to mitigate the crippling effects of overfitting, by introducing random diversity each time a sample is passed through the model and to overcome the weaknesses of the zero-shot tumor classifier, for which we can't measure the performance. Moreover, if the classifier happens to select patches that are not relevant to our task, the information content for that image is lost.

4 Experiments and results

The results we obtained using the siamese architecture for the classification of survival time were measured over an average of 20 independent trainings, to make up for the huge variance introduced by the minuscule dataset we had to deal with. Indeed, the number of cases we could use dropped significantly when we filtered out the non-suitable samples. We only kept the cases that had:

- Methylation data
- Gene expression data
- at least one WSI image
- survival time label

Given all of those constraints, the suitable cases were just 259.
To measure our performance we used the following metrics:

- Accuracy
- Precision
- Recall
- AUC: area under the curve

We performed a great number of tests to find the best mix of inputs and hyperparameters, we tested:

- With or without the standard scaler for omics data
- Apply or not PCA on the omics data
- using FPKM or FPKM-UQ as gene expression

The result in which we got the best performance out of our architecture was achieved using three inputs: Methylation beta values, Gene expression values (FPKM), and WSI embeddings. In particular, we saw a slight enhancement in the performance thanks to the addition of the images with respect to using just the omics. The results are shown in figure 5 and in Table 2. We show a boxplot of the results of 20 trainings, showing how the average values for all the metrics are better with the use of images. In particular, the best setup involved the use of the random patch selector to generate the WSI embeddings, while the performance obtained with the patch classifier was slightly worse, likely due to its low accuracy, and also for the fact that, in any case, the model was not able to 'see' most of the image, leaving important information unused.

We also tried adding copy number variations as a fourth input, but in this case, the metrics dropped, likely because of the high dimensionality of the CNV vectors worsening the problem of overfitting.
In our best results, we trained for 80 epochs using an ADAM optimizer. The learning rate was set to 0.001 and halved after 50 epochs. The Siamese feature extraction network was set to be 3 layers deep and so was the classifier MLP. Another important detail is that in the original architecture [6] the classifier

	Ours						Customics [5]	Siamese [6]
WSI patches	✓	✓	✓			✓		
Gene Expression	✓	✓	✓	✓	✓		✓	✓
Methylation	✓	✓		✓	✓	✓	✓	✓
CNV		✓			✓		✓	
MiRNA								✓
Accuracy	**0.69**	0.62	0.67	0.65	0.64	0.64	0.67	0.64

Table 2: Results of our architecture and comparison with the other solution found in the literature. We replicated their experiments with our data, noticing how ovarian cancer is a much more difficult cohort than others since the accuracy for this type of tumor is worse than the average, and we observed a lack of comparable findings in existing literature comparisons.

	GBM	BIC	KRCCC	LSCC	OVCA
Accuracy	0.86	0.62	0.79	0.66	0.64
AUC	0.83	0.63	0.80	0.71	0.54

Table 3: Results of the siamese network from [6], where GBM is glioblastoma multiforme, BIC is breast invasive carcinoma, KRCCC is kidney renal clear cell carcinoma, LSCC is lung squamous cell carcinoma, OVCA is ovarian cancer.

was trained using both the cross-entropy loss and the contrastive loss, but we decided to use just the cross-entropy, since the contrastive loss value tended to be overwhelmingly high compared to the other, leading to the non-optimization of the classification task. The model learned really well the distribution of the input training set, reaching a train accuracy as high as 0.95, as seen in the figure 6, the features learned from the training set are indeed well representative of our classification task, but this is not true when it comes to generalization, as our performance drops dramatically in the inference scenario. By conducting some ablation studies, we tried to quantify the importance of the three inputs we gave to the model, so, as seen in figure 5, the most significant performance drop comes from the removal of the gene expression data, while the worst results come with the use of images and methylation data, which seems to be the feature set from which we extract the least meaningful information.

As seen in the previous figures 6, the feature extractor learns a really good representation for the training set, but the capabilities of generalization on the test set leave a lot to be desired, due to the scarcity of the data, but also due to the many factors that can influence the survival time of an individual, that are way out of the scope of this project.

Moreover, we observe that, due to the scarcity of available data and the complexity of the disease, OVCA seems to be a difficult case to study, as the performance results are consistently worse if methods are applied to this scenario, as in table 3.

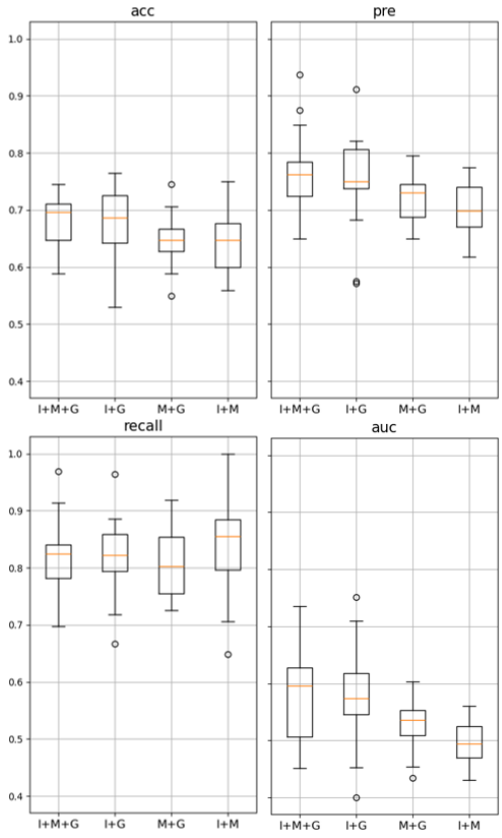
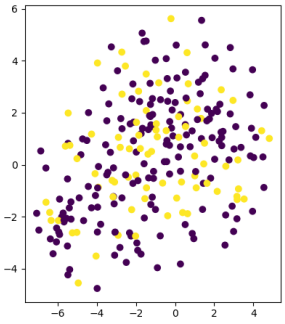


Figure 5: Comparison between the results obtained by using different combinations of the three inputs, I = images, M = Methylation, G = Gene Expreession

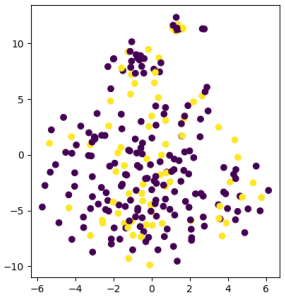
5 Conclusions

In this paper, we explored two alternative methods to perform the integration of heterogeneous genetic data and WSI images in order to improve the performance on the prediction of the overall survival of patients affected by ovarian cancer. The results we obtained are comparable with other state-of-the-art methods for survival time prediction, but they are heavily affected by severe problems related to the scarcity of data and abundance and noisiness of the features.

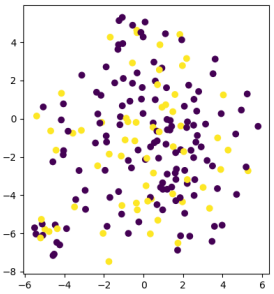
The advantage of our architecture with respect to the others available lies in our unique approach to integrating image and molecular data in a lightweight Siamese



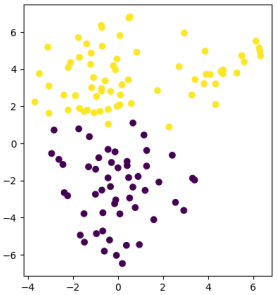
(a) TSNE of the input Gene Expression. The two classes are represented with yellow and purple



(b) TSNE of the input Methylation



(c) TSNE of our input vectors (all inputs concatenated)



(d) TSNE of the feature vectors extracted from the input. This is the entire training set.

Figure 6: TSNE before and after the training

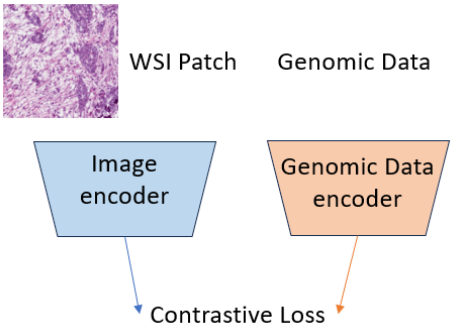


Figure 7: Architecture for image-omics correlation, not implemented yet

network, an addition that gives us a measurable increase in performance, despite not relying on complex end-to-end pretrained architectures [12]. Moreover, the proposed architecture is easily configurable and adaptable to process different types of input data, also varying the number of modalities.

Future developments of this architecture could focus on better feature selection methods for both the omics and the images as well as the development of an end-to-end network that would overcome the lack of a MIL-based patch selection mechanism.

Moreover, the model could take into account the general clinical data, leveraging on information coming from background knowledge of the patient, such as age and ethnicity, however, we chose to avoid this route because it would have likely introduced shortcuts in our data, leading to a heavily biased model. We could also extend our project by introducing a self-supervised component that would solve the absence of labels during the tumoral patch selection and better learn the features of the multi-omics. An example of this approach can be seen in [18].

Another task that could be explored is the correlation between omics data and WSIs, in order to find the parts of the images that more closely resemble the genetic characteristics of the depicted tumor. We propose a simple architecture to perform this task, inspired by the many successful use cases of contrastive learning in order to learn a shared representation space. In particular, we have two independent encoders, jointly trained thanks to a contrastive loss. Then, at inference time, given the genetic data and WSI image of a patient, we would simply need to divide the image into patches and compute a similarity score between the embeddings of the patch and the omics, to hopefully find areas of the WSI image that are more relevant to the genes, and vice versa.

In conclusion, models leveraging transformer attention and early fusion techniques appear to be promising solutions for addressing the challenges in our study, particularly when there is a significant amount of data available for analysis [12].

References

[1] cancerresearchuk.org, "Ovarian cancer statistics - cancerresearchuk.org." https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/ovarian-cancer, 2019. [Accessed 31-01-2024].

[2] S. R. Stahlschmidt, B. Ulfenborg, and J. Synnergren, "Multimodal deep learning for biomedical data fusion: a review," *Briefings in Bioinformatics*, vol. 23, p. bbab569, 01 2022.

[3] B. Mirza, W. Wang, J. Wang, H. Choi, N. C. Chung, and P. Ping, "Machine learning and integrative analysis of biomedical big data," *Genes*, vol. 10, p. 87, Jan. 2019.

[4] W. Qiu, J. Yang, B. Wang, M. Yang, G. Tian, P. Wang, and J. Yang, "Evaluating the microsatellite instability of colorectal cancer based on multimodal deep learning integrating histopathological and molecular data," *Frontiers in Oncology*, vol. 12, 2022.

[5] H. Benkirane, Y. Pradat, S. Michiels, and P.-H. Cournède, "Customics: A versatile deep-learning based strategy for multi-omics integration," *PLOS*

Computational Biology, vol. 19, p. e1010921, Mar. 2023.

[6] Y. Hao, X.-Y. Jing, and Q. Sun, "Cancer survival prediction by learning comprehensive deep feature representation for multiple types of genetic data," *BMC Bioinformatics*, vol. 24, p. 267, Jun 2023.

[7] M. Y. Lu, D. F. K. Williamson, T. Y. Chen, R. J. Chen, M. Barbieri, and F. Mahmood, "Data efficient and weakly supervised computational pathology on whole slide images," 2020.

[8] Z. Huang, F. Bianchi, M. Yuksekgonul, T. J. Montine, and J. Zou, "A visual–language foundation model for pathology image analysis using medical twitter," *Nature Medicine*, pp. 1–10, 2023.

[9] J. Breen, K. Allen, K. Zucker, P. Adusumilli, A. Scarsbrook, G. Hall, N. M. Orsi, and N. Ravikumar, "Artificial intelligence in ovarian cancer histopathology: a systematic review," *npj Precision Oncology*, vol. 7, Aug. 2023.

[10] F. Zhan, L. He, Y. Yu, Q. Chen, Y. Guo, and L. Wang, "A multimodal radiomic machine learning approach to predict the lck expression and clinical prognosis in high-grade serous ovarian cancer," *Scientific Reports*, vol. 13, Sept. 2023.

[11] R. J. Chen, M. Y. Lu, W.-H. Weng, T. Y. Chen, D. F. Williamson, T. Manz, M. Shady, and F. Mahmood, "Multimodal co-attention transformer for survival prediction in gigapixel whole slide images," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 4015–4025, October 2021.

[12] G. Jaume, A. Vaidya, R. Chen, D. Williamson, P. Liang, and F. Mahmood, "Modeling dense multimodal interactions between biological pathways and histology for survival prediction," 2023.

[13] R. L. Allesøe, A. T. Lundgaard, R. Hernández Medina, and e. a. Aguayo-Orozco, Alejandro, "Discovery of drug–omics associations in type 2 diabetes with generative deep-learning models," *Nature Biotechnology*, vol. 41, p. 399–408, Jan. 2023.

[14] M. T. Hira, M. A. Razzaque, C. Angione, J. Scrivens, S. Sawan, and M. Sarker, "Integrated multi-omics analysis of ovarian cancer using variational autoencoders," *Scientific Reports*, vol. 11, Mar. 2021.

[15] H. Uno, T. Cai, M. Pencina, R. D'Agostino, and L. Wei, "On the c-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data," *Statistics in medicine*, vol. 30, pp. 1105–17, 05 2011.

[16] A. Gretton, K. M. Borgwardt, M. J. Rasch, B. Schölkopf, and A. Smola, "A kernel two-sample test," *Journal of Machine Learning Research*, vol. 13, no. 25, pp. 723–773, 2012.

[17] J. L. Katzman, U. Shaham, A. Cloninger, J. Bates, T. Jiang, and Y. Kluger, "Deepsurv: personalized treatment recommender system using a cox proportional hazards deep neural network," *BMC medical research methodology*, vol. 18, no. 1, pp. 1–12, 2018.

[18] S. Hashim, K. Nandakumar, and M. Yaqub, "Self-omics: A self-supervised learning framework for multi-omics cancer data," 2022.