# Preparation of Papers for IEEE TRANSACTIONS ON MEDICAL IMAGING

First A. Author, *Fellow, IEEE*, Second B. Author,and Third C. Author, Jr., *Member, IEEE*

***Abstract*—**

***Index Terms*— Enter about five key words or phrases in alphabetical order, separated by commas. Cancer survival analysis,Cancer survival prediction,deep learning,whole slide image**

## I. Introduction

Survival prediction is a direction of statistics for analyzing the duration time that is expected until the events of interest occur,such as the death of the life form of biology. With the rising importance of precision medicine, the ability to predict patient survival based on individualized characteristics becomes paramount. WSI, as a rich source of information, offers an unprecedented level of granularity by capturing the heterogeneity within tissue samples. By enabling the high-resolution digitization of entire tissue slides, WSI has not only enhanced the efficiency and accuracy of pathological assessments but has also opened new avenues for predictive modeling and survival analysis.

This review aims to provide a comprehensive overview of the application of WSI in the context of survival prediction. We delve into the intersection of digital pathology and computational methods, exploring how WSI data can be harnessed to predict patient outcomes, especially in the field of oncology.

The introduction of WSI has given researchers and clinicians access to vast repositories of image data, offering new opportunities for feature extraction, machine learning, and deep learning techniques to model and predict patient survival. The ability to analyze tissue morphology, cellular structures, and patterns within whole slide images has the potential to uncover valuable insights into disease progression and prognosis.

In this review, we will examine the key methodologies, challenges, and recent advancements in survival prediction using WSI data. We will also discuss the implications of such predictions on personalized medicine, treatment strategies, and patient care.Additionally, we will explore various aspects of survival prediction using WSI, including the development of novel image features, the utilization of deep learning architectures for image analysis, and the challenges associated with large-scale data management. We will also consider the ethical and regulatory considerations in implementing such predictive models in clinical practice.

As the fields of digital pathology and computational biology continue to evolve, the integration of WSI into survival analysis represents a promising frontier for improving patient outcomes and advancing our understanding of disease dynamics. This review will serve as a guide for researchers and practitioners interested in leveraging WSI for survival prediction in diverse clinical and research settings.

## II. Background

In recent years, several computer algorithms for hematoxylin and eosin (H&E) stained pathology image analysis have been developed to aid pathologists in objective clinical diagnosis and prognosis. The Cancer Genome Atlas (TCGA) are the most common datasets when discussing the topics of survival prediction.

the observation of one patient is either a survival time $(O_i)$ or a censored time $(C_i)$. If and only if $t_i = \min(O_i, C_i)$ can be observed during the study, the dataset is right- censored [17]. An instance in the survival data is usually represented as $(x_i, t_i, \delta_i)$ where $x_i$ is the feature vector, $t_i$ is the observed time, $\delta_i$ is the indicator which is 1 for a un- censored instance (death occurs during the study) and 0 for a censored instance.

The survival function $S(t|x) = Pr(O \geq t|x)$ is used to identify the probability of being still alive at time $t$ where $x = (x_1, ...x_p)^T$ is the covariates of dimension $p$, The hazard function is defined as

In recent years, the integration of Whole Slide Imaging (WSI) into the realm of survival prediction has emerged as a promising frontier in the fields of digital pathology, oncology, and computational biology. WSI, a revolutionary technology, has transformed the way we analyze tissue samples, providing a comprehensive digital representation of entire pathology slides.

Traditionally, pathological assessments have been reliant on manual inspection and subjective interpretation of glass slides under a microscope. While this approach has served as the gold standard for disease diagnosis and prognosis, it is

This paragraph of the first footnote will contain the date on which you submitted your paper for review. It will also contain support information, including sponsor and financial support acknowledgment. For example, "This work was supported in part by the U.S. Department of Commerce under Grant BS123456."

The next few paragraphs should contain the authors' current affiliations, including current address and e-mail. For example, F. A. Author is with the National Institute of Standards and Technology, Boulder, CO 80305 USA (e-mail:author@boulder.nist.gov).

S. B. Author, Jr., was with Rice University, Houston, TX 77005 USA. He is now with the Department of Physics, Colorado State University, Fort Collins, CO 80523 USA (e-mail: author@lamar.colostate.edu).

T. C. Author is with the Electrical Engineering Department, University of Colorado, Boulder, CO 80309 USA, on leave from the National Research Institute for Metals, Tsukuba, Japan (e-mail: author@nrim.go.jp).

inherently limited by issues of inter-observer variability, labor intensiveness, and the inability to harness the full potential of the vast data embedded in tissue structures.

The advent of WSI has paved the way for a paradigm shift in tissue analysis. Through high-resolution digitization, WSI generates massive datasets of tissue samples, allowing for a more detailed examination of cellular structures, tissue morphology, and spatial relationships. These digital representations of tissue hold a wealth of information that extends far beyond what can be appreciated by the human eye.

One of the most compelling applications of WSI lies in the realm of survival prediction, particularly in the context of cancer. Cancer remains a leading cause of mortality worldwide, and the ability to predict patient outcomes accurately is paramount for optimizing treatment strategies and improving patient care. With the integration of WSI, researchers and clinicians can extract rich image features and employ advanced machine learning and deep learning techniques to model and predict patient survival.

Survival prediction using WSI offers the potential to uncover subtle patterns, biomarkers, and prognostic factors that may have gone unnoticed in traditional pathology assessments. The granularity and comprehensiveness of WSI data open doors to new avenues of research and personalized medicine, enabling clinicians to tailor treatments to individual patient profiles.

In this review, we delve into the intricate interplay between digital pathology and computational methods, exploring how WSI is being harnessed to predict patient survival. We will examine the methodologies, challenges, ethical considerations, and the implications of these predictions on clinical practice.

As the field of WSI continues to evolve, the fusion of digital pathology and survival prediction holds great promise for advancing our understanding of disease dynamics and ultimately enhancing patient outcomes.

As we delve deeper into this evolving landscape, it is crucial to appreciate the monumental shift brought about by WSI. Its capacity to create digital archives of pathology slides has not only expedited the diagnostic process but has also catapulted computational pathology to the forefront. This dynamic shift, driven by advancements in digital imaging and artificial intelligence, has redefined the scope of pathology by offering a more profound understanding of disease, bolstering diagnostic accuracy, and heralding a new era of prognostic modeling.

WSI's role in survival prediction is particularly significant within the domain of cancer research. It has given rise to the emergence of predictive models that harness the extensive image data contained within tissue samples. By analyzing and extracting pertinent information from these digitized slides, researchers can identify morphological and structural biomarkers, which, when combined with clinical data, offer valuable insights into patient outcomes.

Moreover, the integration of WSI data provides the foundation for the development of prognostic models capable of guiding treatment decisions and improving the overall quality of patient care. The implications extend beyond individualized medicine to population-level studies, ultimately influencing healthcare policy, resource allocation, and public health strategies.

While the potential is immense, it is not without challenges. The management and analysis of vast WSI datasets demand advanced computational infrastructure, robust machine learning algorithms, and a harmonious partnership between pathologists and data scientists. Ethical considerations surrounding patient data privacy, model interpretability, and regulatory compliance also merit attention.

This review embarks on a journey to explore the multifaceted landscape of survival prediction using WSI. It endeavors to survey the methodologies, showcase the milestones achieved, and address the formidable challenges that lie ahead. In doing so, it aspires to serve as a guiding compass for researchers, pathologists, clinicians, and healthcare stakeholders invested in leveraging the symbiotic relationship between digital pathology and survival prediction.

As we advance in this ever-evolving field, the fusion of digital pathology with survival prediction promises to unravel the intricacies of disease dynamics, enhance patient care, and pave the way for a new era of data-driven medicine.

### A. Abbreviations and Acronyms

Define abbreviations and acronyms the first time they are used in the text, even after they have already been defined in the abstract. Abbreviations such as IEEE, SI, ac, and dc do not have to be defined. Abbreviations that incorporate periods should not have spaces: write "C.N.R.S.," not "C. N. R. S." Do not use abbreviations in the title unless they are unavoidable (for example, "IEEE" in the title of this article).

## III. Methodology

### A. Region of Interest

There exist many methods developed for predicting survival through the information provided by the whole slide images(WSIs). Rather than utilizing the overall patches from the gigapixel pathology images, the traditional models usually preselect a subset of critical patches from the region of interest (ROI) as the input data. Apart from the features extracted from ROIs with the deep neural networks(DNNs), some morphological features of the image patches can be extracted and accessed by the image analysis software named CellProfiler [1], commonly used for cell phenotypes measurements at that time. It's the features including cell shape, size, the distribution of pixel intensity in the cells and nuclei and texture of cells and nuclei that the quantitative analysis tool can extract. Taking whole slide images with various size as inputs, in order to do end-to-end survival prediction, DeepConvSurv proposed by [2] randomly chooses the patches among the ROIs annotated by the professional pathologists. And the experiments showed it could extract more abstract information different from the hand-crafted features generated by the state-of-the-art analysis tool CellProfiler mentioned above. Except this, [3], [4], [5], [6], [7] have shown that random sampling of patches within the tissues in WSIs still makes sense in stratifying phenotypic information which can be improved. With the help of online tool, Mobadersany et al. [8] manually selected the ROIs

without tissue-processing artifacts containing overstaining or understaining areas from alternates, which have viable tumor features. Faced the difficulty of intratumoral heterogeneity and few availability of labeled data, to obtain better and more robust effects, the model uses the data augmentation techniques and sorted median risks to get prediction results. And the model eventually gets more accurate outcomes. Moreover, the paper provides some publicly accessible datasets with ROIs. With the assistance of the assembled diagnostic slides offered by [8] with ROIs, characteristics of tumor microenvironment could be got in the built graphs in [9]. Because semantic segmentation is executed in ROIs to recognize and localize relevant cells acting as set of nodes in the spatial graph for abstract graph representations. Not only dose the model in [10] use the public cancer survival dataset TCGA, but it also adopts a core sample set from UT MD Anderson Cancer Center during the holistic procedure. And it takes advantage of the annotations of ROIs to locate the possible tumor regions in pathological images for subsequent steps. Some methods adopt sampling strategy to generate candidate patches not limited to ROIs. Since revealing ROIs requires specialized prior knowledge and expensive labor costs, Wang et al. proposed an automatic model aimed at finding ROIs. The proposed model in [11] can identify tumor regions as ROIs in hematoxylin and eosin (H&E) stained pathology images using predicted likelihood of image patches, each patch tagged as the highest probability category. In this way, some tumor area-related features can be extracted as the descriptors of above ROIs including area, perimeter, convex area, filled area, major axis length, minor axis length and so on. For the purpose of training the prognostic model indicating that the risk group defined by tumor shape features is an independent prognostic factor, the features are used in the training process. The model in [12] has a automated pipeline excluding the background white space among the identified tissue, then overlooking the sparse cellularity regions and randomly sampling the potential patches from the foreground area.

In a short, as the methods previously implemented, ROIs routinely ask artificial marking and rigorous reviews to produce passable survival prediction. Additionally speaking, the ROI-based methods discussed above require pathologists to hand-annotate ROIs, a tedious task.

### B. Feature Extraction

Previous methods often extract image features from patches of whole slide images(WSIs) using the pre-trained model based on the numerous natural images datasets ImageNet theoretically being able to withdraw the low-dimensional features such as edge and texture.However,they have ignored the enormous difference between the WSIs and the natural images. Recently, some new methods have been proposed to suitably get features from the WSIs to overcome the significant shortcomings. Without the knowledge of each patch-level labels, the self-supervised learning methods can autonomously impart the outstanding feature extraction ability to the model.

The method named SimCLR in [13], one of the self-supervised learning(SSL) methods, using the contrastive learning has excellent feature extraction capacity even comparable to the supervised learning model thanks to the collected 57 digital histopathology datasets with none labels. SeTranSurv, the proposed model in [14] applies SimCLR to train the initial feature extraction model ResNet18 [15] to get the specialized model for our downstream task survival prediction. During the training the model applies the contrastive loss [16] to enhance feature extraction ability. Firstly augmentation module in SeTranSurv using two methods to transform a random image example to a position pair, then the model uses the fine-tuned feature extractor raised above to get the features. Considering other images as negative examples ensures that the views of different slide images are far apart in the high-dimensional space and the views of the same image are closer in the process of training. Additionally, position encodings capturing spatial information and self-attention modules learning correlation between patches are put into the training of the model above to obtain slide-level features and therefore the patient-level features. Aside from SimCLR as a strategy used in [17] and [18], KimiaNet from [19] has also been one of the most welcome pre-trained models used for survival analysis exploiting a variable, multi-organ open image repository lick TCGA, which has been employed directly to extract the feature embeddings of image patches in [20] [18]. Without using convnet-based methods, the model in [21] also considers the self-supervised learning (SSL) methods mainly for making full use of plentiful color information designed for WSI patches or pixels without hand-actuated labels, totally colorization and cross-channel as the pretext tasks. The colorization model is trained to predict corresponding color channels based on the lightness channel and the letter, on the other hand, is trained to get lightness channel using given color channels data, after which the visualized results indicating highlighted overall structure of nuclei and tissue.

### C. Multimodal

One limitation for some existing survival models is that they initially focus on one modality and cannot sufficiently handle multi-modalities data. Actually, multi-modalities information could provide complementary and auxiliary information for tumor diagnosis. For instance, molecular data and whole slide images share relevant representations to describe the same event in tumor growth and symptoms which are very critical for tumor diagnosis. Therefore,it is essential and necessary to combine and integrate multi-modal data such as pathological images, genotypic information and clinical data for explaining and understanding cancerous heterogeneity and complex symptoms for customized treatments and healings, consequently boosting the survival predictions. Although hematoxylin and eosin (H&E)-stained slides are enough to build a comprehensive diagnosis, other modal data can provide a deeper description of the tumor. For example, genomic profiles being comprising of ten thousand dimensional sequences can provide a molecular characterization of the tumor. Additionally,during the past several decades, multiple clinicians have made clinical cancer survival prediction on the basis of clinical variates and experience, therefore the clinical data is also an important source modality for multi-modal survival

prediction. However, multi-modal survival prediction faces an important challenge due to the huge data heterogeneity gap between WSIs and other modalities, and many proposed approaches use simple multi-modal fusion mechanisms for feature incorporation, which give up mining important multi-modal relationships.

The model in [22] gets the features from the pathological images extracted using Cellprofiler [1] from the image tiles in ROIs,namely geometry, texture, holistic features.In addition,the model uses the preprocessed genetic data and one feature selection operation SPACE [23] to select representative features to integrate data for the principal component regression model for survival prediction. The experiments focusing on ADC lung cancer revealed that the results could be better than only using data of genes or images, which demonstrated that the genetic data could actually enhance the prediction performance of the survival analysis model to some extent.

Differently, the experiments in [10] were conducted on other two cancer types: glioblastoma multiforme (GBM) and lung squamous cell carcinoma (LUSC) using pathological images and molecular data including protein data, Copy number variation(CNV) data and so on. But similarly, to fuse the data from different modality for better results, the model firstly learn deep representations from two kinds of data using separate Convolutional Neural Networks (CNNs).Next, the representations passing through the sub-network in the model are connected to get the new representation, which serves as the input of the correlational layer. Certainly, the deep correlation layer is used to decrease the discrepancy by maximizing the correlation. After the common layer, the output acts as the input of the survival prediction layer using the negative log partial likelihood as survival loss function. Compared with the models handling the linear condition, the new model DeepCorrSurv can learn complex correlation using deep neural networks by using the unsupervised method to learn the interactions and the survival loss network to fine-tune the model, eventually getting better results in the comparison experiments about lung and brain cancer. The results showed that the common representation after maximizing step could bring better performance to the survival prediction measured by the metric named concordance index values or c-index values.

In the paper [8], the model named GSCNN aims for fusing genomic data and second modal data from The Cancer Genome Atlas (TCGA) Lower-Grade Glioma (LGG) and Glioblastoma (GBM) projects. If the second modal data are added to the network during the whole training process, the median c index will improve more than simply integrating the second type data into the fully connected layers. Then the model proves that the Molecular subtype is significant in the multi-variable regression model.

Again, for genomic information, Sun et al. in [24] considered building distinct models to make survival predictions according to the five major molecular subtypes of the breast cancer. The genomic data chosen in the model consists of gene expression, copy number alteration (CNA), gene methylation and protein expression. The main idea in the paper is how to merge different types of data as one type, and in this way using multiple kernel learning (MKL) is a choice. To integrate the genotypic information and image data, GPMKL model was proposed using 5 independent Gaussian kernels and having a integrating step. The baseline algorithms used that time for comparison includes LASSSO-Cox [25], elastic-net penalized Cox (EN-Cox) [26], Parametric censored regression models (PCRM) [27], Random survival forests (RSF) [28], Boosting concordance index (BoostCI) [29], Supervised principal components regression (superPC) [30]. Additionally, two independent models named GMKL and PMKL was constructed only adopting genomic data or pathological images data and four single dimensional models using four types of genomic data were also built. Subsequently, the results showed that the gene expression and protein information play relatively more important role than others and CNA makes a little contribution to holistic prediction accuracy.

In the model proposed in [31] named MultimodalPrognosis, clinical, genomic and WSI images data are processed by FC layers, deep highway networks [32], the SqueezeNet architecture [33] separately. The microRNA data is also passed through deep highway networks discussed. One notable problem about the microRNA and clinical data is the missing data. However, the function of the highway networks dose not stand out without comparison with other methods in the paper. Then, the similarity loss is used to train the model to recognize the patient-distinguishing patterns and correspond the data from different modality to generate associated representations. Thus, the final loss function is composed of cox loss and similarity loss in the unsupervised model. The multi-modal dropout was also invented that is dropping whole feature vectors of each modality and accordingly increase the weights of other modalities to build robust representations. The solution above was then validated in the experiments including visualizing the encodings of the pancancer patient cohort and calculating the C-index values and the results also demonstrated the essence of the adopted modalities.

One more variation usually ignored is the ordinal relationship among the survival time of different patients, which is assumed independent among patients. The model proposed in [34] named OSCCA intends to take advantage of the information. For extracting features, the model uses the methods in [35] to generate segmented nucleus and get the specific features. As to gene expression data, the co-expression network analysis algorithms are utilized to derive eigen-gene features. Considering the correlation between imaging and eigen-gene features, sparse canonical correlation analysis(i.e. SCCA) is used to choose features. Among the datasets, most patients are censored, denoting that their actual living time is longer than recorded data, while the uncensored patients have the real survival time. To make full use of the censored state, an ordinal sparse canonical correlation analysis (OSCCA) method is proposed. In the newly proposed method, the equation estimates the uncensored information, while linear inequalities restrains the ordered relationship between censored and uncensored data.

After the model above, Shao et al. developed a new model named OMMFS in [36] using CNV data and level-3 DNA Methylation (DME) data additionally. The experiments

showed that the log-rank test [37] had better stratification performance than univariate Cox regression method. Furthermore, the model implements the second feature selection function based on the Generalized Sparse Canonical Correlation Analysis (GSCCA) framework [38] to get the inherent relationship among different modalities. Likewise the modality in OSCCA, the model under discussing also notices the survival information of patients. The validation experiments demonstrated that the new model even had superior stratification of early-stage KIRC patients. To make comparison, the SGSCCA model was proposed with the same objective function without ordered survival information. However, the feature pre-selection strategy is based on the median value, which is too arbitrary to consider the accurate relevance between features and cancer type. Also, the image features relies on the regions of interest annotated by pathologists. This model in [39] has a two-stage feature extraction process used in a deep learning model for pathology-specific layers. In the first stage, a pre-trained convolution neural network (CNN) is used to identify survival-discriminative features in patches of pathological images. These features are obtained through dilated convolution layers and max-pooling. In the second stage, global survival-discriminative features for a whole slide image (WSI) are generated by aggregating feature scores from multiple patches. A two-stage pooling approach, including 3-norm pooling, is used to rank and aggregate the most important features and patches. The resulting vector of aggregated survival-discriminative features represents a WSI for a patient, contributing to the integrative deep learning model. The genome and demography-specific layers in this model are adapted from the Cox-PASNet [40], which is a pathway-based sparse deep neural network. The genome-specific layers consist of a gene layer, a pathway layer, and two hidden layers (H1 and H2). The gene layer serves as an input layer for gene expression data, with each node representing a gene. The pathway layer incorporates prior biological knowledge from databases like KEGG for biological interpretation. Connections between the gene layer and pathway layer are established based on biological pathway databases, with pathway nodes representing specific biological pathways. The two hidden layers capture nonlinear and hierarchical relationships between the pathways. Clinical patient data are integrated into the demography-specific layer and combined with genomic features from gene expressions and aggregated survival-discriminative features from pathological images in the final hidden layer of the integrative model. To address overfitting in deep learning models with high-dimensional, low-sample-size data, the training technique from Cox-PASNet is applied. Instead of training the entire network, small networks are randomly selected, and sparse coding is used to create sparse connections for model interpretability. Training continues until convergence, with validation data used to monitor errors and prevent overfitting through early stopping. PAGE-Net statistically outperformed Cox-EN with histopathological images only and Cox-PASNet with genomic data only.

In the model proposed in [9], the cell graphs from histology images are supposed to get the cell-to-cell interactions and cell neighborhood structure. To build cell graphs, the first step is generating accurate nuclei segmentation, in which a conditional generative adversarial network (cGAN) is used to learn the appropriate loss function for semantic segmentation. The edge set and adjacency matrix of the graph are constructed using the K-Nearest Neighbors (KNN) algorithm from segmented cell nuclei. In addition to manually computed statistics, an unsupervised technique called Contrastive Predictive Coding (CPC) is employed to extract 1024-dimensional features from tissue regions centered around each cell. Graph Convolutional Networks (GCNs) learn abstract feature representations for each node by aggregating feature vectors from their neighborhood through message passing. In scenarios with a high-dimensional feature space and limited training samples, traditional feedforward neural networks are susceptible to overfitting. To address the challenge and apply more robust regularization techniques when training feed-forward networks on high-dimensional, low-sample-size genomics data, the model adopts normalization layers inspired by Self-Normalizing Networks introduced by Klambaeur et al [41]. Moreover, the Kronecker Product is used to construct a multi-modal representation. The feature vectors of histology images, cell graphs, and genomic features undergo matrix outer product operations to create a multi-modal tensor. This tensor captures important interactions among these three modalities in terms of single-modal, bimodal, and tri-modal relationships. Ultimately, a neural network is trained using fully connected layers with the multi-modal tensor as input. The central aim of this method is to fuse heterogeneous modalities with distinct structural dependencies, thereby enhancing research and analysis in cancer pathology. To mitigate the impact of noisy uni-modal features during multi-modal training, a gating-based attention mechanism [42] is introduced to control the expressive power of features within each modality. When fusing histology images, cell graphs, and genomic features, the gating mechanism helps reduce the feature space's size before performing the Kronecker Product calculation.

To integrate the multi-modality data with different weights, the model proposed in [43] uses an asymmetrical Transformer encoder. The main idea to fuse other modality data unevenly is to add the new nodes and edges into the original graphs. Different from the normal self-attention in Transformer, the noisy genomic nodes cannot impact the image features because they do not have the outgoing edges, which can only improve themselves by the influence of the imaging features.

Apart from this, another approach to add the genotypic information is using the biological pathway databases to teach the model about the hidden biological functionality. PONET proposed in [44] uses a sparse biological pathway-informed embedding network for gene expression, additionally adopting the Multi-modal Factorized Bilinear pooling (MFB) method instead of original bilinear model to generate unimodal fusion to catch the modality-specific representations. Getting each uni-modal fusion, the model can use the output representations as the input of the bimodal and tri-modal fusion respectively utilizing the bimodal attention and tri-modal attention. Finally the model is trained through the Cox partial likelihood loss proposed by [31] used for the multi-modalities survival prediction to get the prediction results.

GPDBN in [45] uses Kronecker product to build inter-modality and intra-modality interactions between pathology and genomic data for cancer prognosis prediction.

To overcome the limitation of Kronecker product, HFB-Surv extended GPDBN mentioned with the factorized bilinear model.

### D. WSI Features Fusion

Use "soft" (e.g., `\eqref{Eq}`) cross references instead of "hard" references (e.g., `(1)`). This will make it possible to combine sections, add equations, or change the order of figures or citations without having to manually change equation references.

Do not use the `{eqnarray}` equation environment. Use `{align}` or `{IEEEeqnarray}` instead. The `{eqnarray}` environment leaves unsightly spaces around relation symbols.

Note that the `{subequations}` environment in LaTeX will increment the main equation counter even when there are no equation numbers displayed.

BibTeX only functions in conjunction with local .bib files. If you use BibTeX to produce the bibliography you must attach the .bib files.

LaTeX can't read your mind. If you assign the same label to both a subsubsection and a table, you may find that Table I has been cross referenced as Table IV-B3.

LaTeX does not have precognitive abilities. If you put a `\label` command before the command that updates the counter it's supposed to be using, the label will pick up the last counter to be cross referenced instead. In particular, a `\label` command should not go before the caption of a figure or a table.

Do not use `\nonumber` inside the `{array}` environment. It will not stop equation numbers inside `{array}` and it might stop a wanted equation number in the surrounding equation.

If you are submitting your paper to a colorized journal, you can use the following two lines at the start of the article to ensure its appearance resembles the final copy:

```
\documentclass[journal,twoside,web]{ieeecolor}
\usepackage{Journal_Name}
```

## IV. Units

Use either SI (MKS) or CGS as primary units. (SI units are strongly encouraged.) English units may be used as secondary units (in parentheses). For example, write "1 kg (2.2lb)." An exception exists for when English units are used as identifiers in commercial products, such as a "3½-in disk drive." Avoid combining SI and CGS units, such as current in amperes and magnetic field in oersteds. This often leads to confusion because equations do not balance dimensionally. If you must use mixed units, clearly state the units for each quantity in an equation.

The SI unit for magnetic field strength $H$ is A/m. However, if you wish to use units of T, either refer to magnetic flux density $B$ or magnetic field strength symbolized as $\mu_0 H$. Use the center dot to separate compound units, e.g., "A·m$^2$."

TABLE I
Units for Magnetic Properties

| Symbol | Quantity | Conversion from Gaussian and CGS EMU to SI [a] |
|---|---|---|
| $\Phi$ | magnetic flux | 1 Mx $\rightarrow 10^{-8}$ Wb $= 10^{-8}$ V·s |
| $B$ | magnetic flux density, magnetic induction | 1 G $\rightarrow 10^{-4}$ T $= 10^{-4}$ Wb/m$^2$ |
| $H$ | magnetic field strength | 1 Oe $\rightarrow 10^3/(4\pi)$ A/m |
| $m$ | magnetic moment | 1 erg/G $=$ 1 emu $\rightarrow 10^{-3}$ A·m$^2$ $= 10^{-3}$ J/T |
| $M$ | magnetization | 1 erg/(G·cm$^3$) $=$ 1 emu/cm$^3$ $\rightarrow 10^3$ A/m |
| $4\pi M$ | magnetization | 1 G $\rightarrow 10^3/(4\pi)$ A/m |
| $\sigma$ | specific magnetization | 1 erg/(G·g) $=$ 1 emu/g $\rightarrow$ 1 A·m$^2$/kg |
| $j$ | magnetic dipole moment | 1 erg/G $=$ 1 emu $\rightarrow 4\pi \times 10^{-10}$ Wb·m |
| $J$ | magnetic polarization | 1 erg/(G·cm$^3$) $=$ 1 emu/cm$^3$ $\rightarrow 4\pi \times 10^{-4}$ T |
| $\chi, \kappa$ | susceptibility | 1 $\rightarrow 4\pi$ |
| $\chi_\rho$ | mass susceptibility | 1 cm$^3$/g $\rightarrow 4\pi \times 10^{-3}$ m$^3$/kg |
| $\mu$ | permeability | 1 $\rightarrow 4\pi \times 10^{-7}$ H/m $= 4\pi \times 10^{-7}$ Wb/(A·m) |
| $\mu_r$ | relative permeability | $\mu \rightarrow \mu_r$ |
| $w, W$ | energy density | 1 erg/cm$^3$ $\rightarrow 10^{-1}$ J/m$^3$ |
| $N, D$ | demagnetizing factor | 1 $\rightarrow 1/(4\pi)$ |

Vertical lines are optional in tables. Statements that serve as captions for the entire table do not need footnote letters.

[a]Gaussian units are the same as cg emu for magnetostatics; Mx = maxwell, G = gauss, Oe = oersted; Wb = weber, V = volt, s = second, T = tesla, m = meter, A = ampere, J = joule, kg = kilogram, H = henry.

## V. Discussion and Future Trends

### A. Challenges in histopathology image analysis

The following list outlines the different types of graphics published in IEEE journals. They are categorized based on their construction, and use of color / shades of gray:

*1) Model interpretability:* Figures that are meant to appear in color, or shades of black/gray. Such figures may include photographs, illustrations, multicolor graphs, and flowcharts.

*2) Clinical translation:* Figures that are composed of only black lines and shapes. These figures should have no shades or half-tones of gray, only black and white.

*3) Author photos:* Not allowed for papers in TMI.

*4) Tables:* Data charts which are typically black and white, but sometimes include color.

### B. Multipart figures

Multipart figures are comprised of more than one sub-figure presented together. If a multipart figure is made up of multiple figure types (one part is lineart, and another is grayscale or color) the figure should meet the strictest applicable guidelines.

### C. File Formats For Graphics

Format and save your graphics as one of the following approved file types: PostScript (.PS), Encapsulated PostScript (.EPS), Tagged Image File Format (.TIFF), Portable Document Format (.PDF), Portable Network Graphics (.PNG), or Metapost (.MPS). After the paper is accepted, any included graphics must be submitted alongside the final manuscript files.

### D. Sizing of Graphics

Most charts, graphs, and tables are one column wide (3.5 inches / 88 millimeters) or page wide (7.16 inches / 181 millimeters). The maximum depth of a graphic is 8.5 inches (216 millimeters). When choosing the depth of a graphic, please allow space for a caption. Authors are allowed to size figures between column and page widths, but it is recommended not to size figures less than column width unless necessary.

### E. Resolution

The proper resolution of your figures will depend on the type of figure it is as defined in the "Types of Figures" section. Author photographs, color, and grayscale figures should be at least 300dpi. Lineart, including tables should be a minimum of 600dpi.

### F. Vector Art

While IEEE does accept and even recommends that authors submit artwork in vector format, it is our policy is to rasterize all figures for publication. This is done in order to preserve figures' integrity across multiple computer platforms.

### G. Colorspace

The term colorspace refers to the entire sum of colors that can be represented within a given medium. For our purposes, the three main colorspaces are grayscale, RGB (red/green/blue) and CMYK (cyan/magenta/yellow/black). RGB is generally used with on-screen graphics, whereas CMYK is used for printing purposes.

All color figures should be generated in RGB or CMYK colorspace. Grayscale images should be submitted in grayscale colorspace. Line art may be provided in grayscale OR bitmap colorspace. Note that "bitmap colorspace" and "bitmap file format" are not the same thing. When bitmap colorspace is selected, .TIF/.TIFF are the recommended file formats.

### H. Accepted Fonts Within Figures

When preparing your graphics IEEE suggests that you use of one of the following Open Type fonts: Times New Roman, Helvetica, Arial, Cambria, and Symbol. If you are supplying EPS, PS, or PDF files all fonts must be embedded. Some fonts may only be native to your operating system; without the fonts embedded, parts of the graphic may be distorted or missing.

A safe option when finalizing your figures is to strip out the fonts before you save the files, creating "outline" type. This converts fonts to artwork that will appear uniformly on any screen.

### I. Using Labels Within Figures

*1) Figure Axis labels:* Figure axis labels are often a source of confusion. Use words rather than symbols. As an example, write the quantity "Magnetization," or "Magnetization M," not just "M." Put units in parentheses. Do not label axes only with units. As in Fig. 1, for example, write "Magnetization (A/m)" or "Magnetization (A·m$^{-1}$)," not just "A/m." Do not label axes with a ratio of quantities and units. For example, write "Temperature (K)," not "Temperature/K."

Multipliers can be especially confusing. Write "Magnetization (kA/m)" or "Magnetization ($10^3$ A/m)." Do not write "Magnetization (A/m) $\times$ 1000" because the reader would not know whether the top axis label in Fig. 1 meant 16000 A/m or 0.016 A/m. Figure labels should be legible, approximately 8 to 10 point type.

*2) Subfigure Labels in Multipart Figures and Tables:* Multipart figures should be combined and labeled before final submission. Labels should appear centered below each subfigure in 8 point Times New Roman font in the format of (a) (b) (c).

### J. Referencing a Figure or Table Within Your Paper

When referencing your figures and tables within your paper, use the abbreviation "Fig." even at the beginning of a sentence. Do not abbreviate "Table." Tables should be numbered with Roman numerals.

### K. Submitting Your Graphics

Format your paper with the graphics included within the body of the text as you would expect to see the paper in print. Please do this at each stage of the review, from first submission to final files. For final files only, after the paper has been accepted for publication, figures should also be submitted individually in addition to the manuscript file using one of the approved file formats. Place a figure caption below each figure; place table titles above the tables. Do not include captions or borders in the uploaded figure files.

### L. File Naming

Figures (line artwork or images) should be named starting with the first 5 letters of the corresponding author's last name. The next characters in the filename should be the number that represents the figure's sequential location in the article. For example, in author "Anderson's" paper, the first three figures might be named ander1.tif, ander2.tif, and ander3.ps.

Tables should contain only the body of the table (not the caption) and should be named similarly to figures, except that '.t' is inserted in-between the author's name and the table number. For example, author Anderson's first three tables would be named ander.t1.tif, ander.t2.ps, ander.t3.eps.

Author photographs or biographies are not permitted in IEEE TMI papers.

### M. Checking Your Figures: The IEEE Graphics Analyzer

The IEEE Graphics Analyzer enables authors to pre-screen their graphics for compliance with IEEE Transactions and Journals standards before submission. The online tool, located at http://graphicsqc.ieee.org/, allows authors to upload their graphics in order to check that each file is the correct file format, resolution, size and colorspace; that no fonts are missing or corrupt; that figures are not compiled in layers or have transparency, and that they are named according to the IEEE Transactions and Journals naming convention. At the end of this automated process, authors are provided with a

detailed report on each graphic within the web applet, as well as by email.

For more information on using the Graphics Analyzer or any other graphics related topic, contact the IEEE Graphics Help Desk by e-mail at graphics@ieee.org.

### N. Color Processing/Printing in IEEE Journals

All IEEE Transactions, Journals, and Letters allow an author to publish color figures on IEEE Xplore® at no charge, and automatically convert them to grayscale for print versions. In most journals, figures and tables may alternatively be printed in color if an author chooses to do so. Please note that this service comes at an extra expense to the author. If you intend to have print color graphics, include a note with your final paper indicating which figures or tables you would like to be handled that way, and stating that you are willing to pay the additional fee.

## VI. Experiments

The prognostic accuracy of models was assessed using Monte Carlo cross-validation.We randomly split our cohort into paired training (80%) and testing (20%) sets.

The word "data" is plural, not singular. The subscript for the permeability of vacuum $\mu_0$ is zero, not a lowercase letter "o." Use the word "micrometer" instead of "micron." A graph within a graph is an "inset," not an "insert." The word "alternatively" is preferred to the word "alternately" (unless you really mean something that alternates). Use the word "whereas" instead of "while" (unless you are referring to simultaneous events). Do not use the word "essentially" to mean "approximately" or "effectively." Do not use the word "issue" as a euphemism for "problem." When compositions are not specified, separate chemical symbols by en-dashes; for example, "NiMn" indicates the intermetallic compound $\mathrm{Ni}_{0.5}\mathrm{Mn}_{0.5}$ whereas "Ni–Mn" indicates an alloy of some composition $\mathrm{Ni}_x\mathrm{Mn}_{1-x}$.

Be aware of the different meanings of the homophones "affect" (usually a verb) and "effect" (usually a noun), "complement" and "compliment," "discreet" and "discrete," "principal" (e.g., "principal investigator") and "principle" (e.g., "principle of measurement"). Do not confuse "imply" and "infer."

Prefixes such as "non," "sub," "micro," "multi," and "ultra" are not independent words; they should be joined to the words they modify, usually without a hyphen. There is no period after the "et" in the Latin abbreviation "*et al.*" (it is also italicized). The abbreviation "i.e.," means "that is," and the abbreviation "e.g.," means "for example" (these abbreviations are not italicized).

A general IEEE styleguide is available at http://www.ieee.org/web/publications/authors/transjnl/index.html.

## VII. Conclusion

A conclusion section is not required. Although a conclusion may review the main points of the paper, do not replicate the abstract as the conclusion. A conclusion might elaborate on the importance of the work or suggest applications and extensions.

## Appendix and the Use of Supplemental Files

Appendices, if needed, appear before the acknowledgment. If an appendix is not critical to the main message of the manuscript and is included only for thoroughness or for reader reference, then consider submitting appendices as supplemental materials. Supplementary files are available to readers through IEEE *Xplore*® at no additional cost to the authors but they do not appear in print versions. Supplementary files must be uploaded in ScholarOne as supporting documents, but for accepted papers they should be uploaded as Multimedia documents. Refer readers to the supplementary files where appropriate within the manuscript text using footnotes. [1]

## Acknowledgment

The preferred spelling of the word "acknowledgment" in American English is without an "e" after the "g." Use the singular heading even if you have many acknowledgments. Avoid expressions such as "One of us (S.B.A.) would like to thank . . . ." Instead, write "F. A. Author thanks . . . ." In most cases, sponsor and financial support acknowledgments are placed in the unnumbered footnote on the first page, not here.

## References

[1] Michael R Lamprecht, David M Sabatini, and Anne E Carpenter. Cellprofiler™: free, versatile software for automated biological image analysis. *biotechniques*, 42(1):71–75, 2007.

[2] Xinliang Zhu, Jiawen Yao, and Junzhou Huang. Deep convolutional neural network for survival analysis with pathological images. In *2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 544–547. IEEE, 2016.

[3] Xinliang Zhu, Jiawen Yao, Feiyun Zhu, and Junzhou Huang. Wsisa: Making survival prediction from whole slide histopathological images. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 7234–7242, 2017.

[4] Donglin Di, Shengrui Li, Jun Zhang, and Yue Gao. Ranking-based survival prediction on histopathological whole-slide images. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 428–438. Springer, 2020.

[5] Jiawen Yao, Xinliang Zhu, Jitendra Jonnagaddala, Nicholas Hawkins, and Junzhou Huang. Whole slide images based cancer survival prediction using attention guided deep multiple instance learning networks. *Medical Image Analysis*, 65:101789, 2020.

[6] Christian Abbet, Inti Zlobec, Behzad Bozorgtabar, and Jean-Philippe Thiran. Divide-and-rule: self-supervised learning for survival analysis in colorectal cancer. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part V 23*, pages 480–489. Springer, 2020.

[7] Jiawen Yao, Xinliang Zhu, and Junzhou Huang. Deep multi-instance learning for survival prediction from whole slide images. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part I 22*, pages 496–504. Springer, 2019.

[8] Pooya Mobadersany, Safoora Yousefi, Mohamed Amgad, David A Gutman, Jill S Barnholtz-Sloan, José E Velázquez Vega, Daniel J Brat, and Lee AD Cooper. Predicting cancer outcomes from histology and genomics using convolutional networks. *Proceedings of the National Academy of Sciences*, 115(13):E2970–E2979, 2018.

[9] Richard J Chen, Ming Y Lu, Jingwen Wang, Drew FK Williamson, Scott J Rodig, Neal I Lindeman, and Faisal Mahmood. Pathomic fusion: an integrated framework for fusing histopathology and genomic features for cancer diagnosis and prognosis. *IEEE Transactions on Medical Imaging*, 41(4):757–770, 2020.

[1] Supplementary materials are available in the supporting documents/multimedia tab. Further instructions on footnote usage are in the Footnotes section on the next page.

[10] Jiawen Yao, Xinliang Zhu, Feiyun Zhu, and Junzhou Huang. Deep correlational learning for survival prediction from multi-modality data. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 406–414. Springer, 2017.

[11] Shidan Wang, Alyssa Chen, Lin Yang, Ling Cai, Yang Xie, Junya Fujimoto, Adi Gazdar, and Guanghua Xiao. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. *Scientific reports*, 8(1):10393, 2018.

[12] Callum Christopher Mackenzie, Muhammad Dawood, Simon Graham, Mark Eastwood, et al. Neural graph modelling of whole slide images for survival ranking. In *Learning on Graphs Conference*, pages 48–1. PMLR, 2022.

[13] Ozan Ciga, Tony Xu, and Anne Louise Martel. Self supervised contrastive learning for digital histopathology. *Machine Learning with Applications*, 7:100198, 2022.

[14] Ziwang Huang, Hua Chai, Ruoqi Wang, Haitao Wang, Yuedong Yang, and Hejun Wu. Integration of patch features through self-supervised learning and transformer for survival analysis on whole slide images. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2021: 24th International Conference, Strasbourg, France, September 27–October 1, 2021, Proceedings, Part VIII 24*, pages 561–570. Springer, 2021.

[15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 770–778, 2016.

[16] Kaiming He, Haoqi Fan, Yuxin Wu, Saining Xie, and Ross Girshick. Momentum contrast for unsupervised visual representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9729–9738, 2020.

[17] Hakim Benkirane, Maria Vakalopoulou, Stergios Christodoulidis, Ingrid-Judith Garberis, Stefan Michiels, and Paul-Henry Cournède. Hyper-adac: Adaptive clustering-based hypergraph representation of whole slide images for survival analysis. In *Machine Learning for Health*, pages 405–418. PMLR, 2022.

[18] Bin Li, Yin Li, and Kevin W Eliceiri. Dual-stream multiple instance learning network for whole slide image classification with self-supervised contrastive learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 14318–14328, 2021.

[19] Abtin Riasatian, Morteza Babaie, Danial Maleki, Shivam Kalra, Mojtaba Valipour, Sobhan Hemati, Manit Zaveri, Amir Safarpoor, Sobhan Shafiei, Mehdi Afshari, et al. Fine-tuning and training of densenet for histopathology image representation using tcga diagnostic slides. *Medical Image Analysis*, 70:102032, 2021.

[20] Ziyu Guo, Weiqin Zhao, Shujun Wang, and Lequan Yu. Higt: Hierarchical interaction graph-transformer for whole slide image analysis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 755–764. Springer, 2023.

[21] Lei Fan, Arcot Sowmya, Erik Meijering, and Yang Song. Cancer survival prediction from whole slide images with self-supervised learning and slide consistency. *IEEE Transactions on Medical Imaging*, 2022.

[22] Xinliang Zhu, Jiawen Yao, Xin Luo, Guanghua Xiao, Yang Xie, Adi Gazdar, and Junzhou Huang. Lung cancer survival prediction from pathological images and genetic data—an integration study. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, pages 1173–1176. IEEE, 2016.

[23] Jie Peng, Pei Wang, Nengfeng Zhou, and Ji Zhu. Partial correlation estimation by joint sparse regression models. *Journal of the American Statistical Association*, 104(486):735–746, 2009.

[24] Dongdong Sun, Ao Li, Bo Tang, and Minghui Wang. Integrating genomic data and pathological images to effectively predict breast cancer clinical outcome. *Computer methods and programs in biomedicine*, 161:45–53, 2018.

[25] Robert Tibshirani. The lasso method for variable selection in the cox model. *Statistics in medicine*, 16(4):385–395, 1997.

[26] Yi Yang and Hui Zou. A cocktail algorithm for solving the elastic net penalized cox's regression in high dimensions. *Statistics and its Interface*, 6(2):167–173, 2013.

[27] John D Kalbfleisch and Ross L Prentice. *The statistical analysis of failure time data*. John Wiley & Sons, 2011.

[28] Hemant Ishwaran, Udaya B Kogalur, Eugene H Blackstone, and Michael S Lauer. Random survival forests. 2008.

[29] Andreas Mayr and Matthias Schmid. Boosting the concordance index for survival data–a unified framework to derive and evaluate biomarker combinations. *PloS one*, 9(1):e84483, 2014.

[30] Eric Bair, Trevor Hastie, Debashis Paul, and Robert Tibshirani. Prediction by supervised principal components. *Journal of the American Statistical Association*, 101(473):119–137, 2006.

[31] Anika Cheerla and Olivier Gevaert. Deep learning with multimodal representation for pancancer prognosis prediction. *Bioinformatics*, 35(14):i446–i454, 2019.

[32] Rupesh Kumar Srivastava, Klaus Greff, and Jürgen Schmidhuber. Highway networks. *arXiv preprint arXiv:1505.00387*, 2015.

[33] Forrest N Iandola, Song Han, Matthew W Moskewicz, Khalid Ashraf, William J Dally, and Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and¡ 0.5 mb model size. *arXiv preprint arXiv:1602.07360*, 2016.

[34] Wei Shao, Jun Cheng, Liang Sun, Zhi Han, Qianjin Feng, Daoqiang Zhang, and Kun Huang. Ordinal multi-modal feature selection for survival analysis of early-stage renal cancer. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 648–656. Springer, 2018.

[35] Hady Ahmady Phoulady, Dmitry B Goldgof, Lawrence O Hall, and Peter R Mouton. Nucleus segmentation in histology images with hierarchical multilevel thresholding. In *Medical Imaging 2016: Digital Pathology*, volume 9791, pages 280–285. SPIE, 2016.

[36] Wei Shao, Zhi Han, Jun Cheng, Liang Cheng, Tongxin Wang, Liang Sun, Zixiao Lu, Jie Zhang, Daoqiang Zhang, and Kun Huang. Integrative analysis of pathological images and multi-dimensional genomic data for early-stage cancer prognosis. *IEEE transactions on medical imaging*, 39(1):99–110, 2019.

[37] Jun Cheng, Jie Zhang, Yatong Han, Xusheng Wang, Xiufen Ye, Yuebo Meng, Anil Parwani, Zhi Han, Qianjin Feng, and Kun Huang. Integrative analysis of histopathological images and genomic data predicts clear cell renal cell carcinoma prognosis. *Cancer research*, 77(21):e91–e100, 2017.

[38] Daniela M Witten and Robert J Tibshirani. Extensions of sparse canonical correlation analysis with applications to genomic data. *Statistical applications in genetics and molecular biology*, 8(1), 2009.

[39] Jie Hao, Sai Chandra Kosaraju, Nelson Zange Tsaku, Dae Hyun Song, and Mingon Kang. Page-net: interpretable and integrative deep learning for survival analysis using histopathological images and genomic data. In *Pacific Symposium on Biocomputing 2020*, pages 355–366. World Scientific, 2019.

[40] Jie Hao, Youngsoon Kim, Tejaswini Mallavarapu, Jung Hun Oh, and Mingon Kang. Cox-pasnet: pathway-based sparse deep neural network for survival analysis. In *2018 IEEE international conference on bioinformatics and biomedicine (BIBM)*, pages 381–386. IEEE, 2018.

[41] Günter Klambauer, Thomas Unterthiner, Andreas Mayr, and Sepp Hochreiter. Self-normalizing neural networks. *Advances in neural information processing systems*, 30, 2017.

[42] John Arevalo, Thamar Solorio, Manuel Montes-y Gómez, and Fabio A González. Gated multimodal units for information fusion. *arXiv preprint arXiv:1702.01992*, 2017.

[43] Ruoqi Wang, Ziwang Huang, Haitao Wang, and Hejun Wu. Ammasurv: asymmetrical multi-modal attention for accurate survival analysis with whole slide images and gene expression data. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 757–760. IEEE, 2021.

[44] Lin Qiu, Aminollah Khormali, and Kai Liu. Deep biological pathway informed pathology-genomic multimodal survival prediction. *arXiv preprint arXiv:2301.02383*, 2023.

[45] Zhiqin Wang, Ruiqing Li, Minghui Wang, and Ao Li. Gpdbn: deep bilinear network integrating both genomic data and pathological images for breast cancer prognosis prediction. *Bioinformatics*, 37(18):2963–2970, 2021.