RESEARCH PAPER

Biological Prior-Driven Dual Graph Attention Network for Concurrent Her2 Amplification and Ordinal Ihc Scoring In Esophageal Adenocarcinoma

Mohod, A. P. [1*,2], Yadav, K. P. [3]

[1]*Department of Computer Science and Engineering MATS University Raipur, CG, India*
[2]*Department of Computer Science and Engineering Priyadarshini Bhagwati Collge of Engineering Nagpur, Maharashtra, India*
[3]*Department of Computer Science and Engineering, MATS University Raipur, CG, India*
Email Id: 2mohod.ashish@gmail.com and 3drkpyadav732@gmail.com
[*]***Corresponding Author:*** *mohod.ashish@gmail.com*
Orcid Id: 20009-0008-0170-5199

Received: 24th April, 2026; Revised: 20th May 2026; Accepted: 30th June, 2026; Available Online: 10th July, 2026

ABSTRACT

Background and Objective: The overexpression of human Epidermal Growth Factor Receptor 2 (HER2) is a key biomarker in esophageal adenocarcinoma. There is a significant inter-observer variability in equivocal IHC 2+ category due to the subjective judgment of clinical category by standard immunohistochemistry (IHC). In fact, current Multiple Instance Learning (MIL) systems cannot handle this ambiguity, because they require patches to be independent and identically distributed, leading to the loss of the information of the tissue topology, and the heavy mixing of morphological and genomic information in the shared embedding multi-tasking.

Methods: We introduce the Biological Prior-Driven Dual Graph Transformer (Bio-DGT) that explicitly forms the biological causal graph between ERBB2 gene amplification and protein phenotype expression. Bio-DGT consists of two Dual-Stream Graph Attention Networks (GATs) to prevent feature entanglement and maintain intercellular topology. A Cross Task Biological Attention (CTBA) module for global genomic inference for active guidance of local morphological graph nodes. At the same time, an Uncertainty Weighted Consistent Rank Logits (UW-CORAL) loss is applied to optimize the discrete IHC task, where the epistemic attention entropy is used to dynamically penalize the predictions in high-ambiguity boundaries.

Results: The independent, strictly stratified esophageal cancer cohort had a binary HER2 accuracy of 0.9790 (AUC: 0.9940) and a multi-class ordinal IHC accuracy of 0.9545 (weighted F1: 0.9551). Importantly, all remaining classification errors were clustered in neighbouring diagnostic classes reflecting the strong constraints imposed by the mathematical order. Spatial interpretability analysis showed that the latent attention topographies are in perfect continuity with the clinical guidelines from the zero/null baseline activations to high topological saturation.

Conclusion: Bio-DGT combines features from spatial tissue topology, biological routing, and uncertainty-calibrated ordinal regression to overcome the problems of cross task feature entanglement. The architecture shifts the focus in computational methods from the ‘intensity of the signal' to the ‘transparency of the morphological reasoning', and provides a biologically plausible and clinically safe decision-support system

Keywords: Esophageal Adenocarcinoma; Immunohistochemistry (IHC); HER2 Amplification; Computational Pathology; Multiple Instance Learning (MIL); Graph Neural Networks (GNN); Ordinal Regression; Multi-Task Learning.

How to cite this article: Mohod AP, Yadav KP, Biological Prior-Driven Dual Graph Attention Network for Concurrent Her2 Amplification and Ordinal Ihc Scoring In Esophageal Adenocarcinoma. Int J Drug Deliv Technol. 2026;16(7): 178-191. DOI: 10.25258/ijddt.16.7.21
Source of support: Nil.
Conflict of interest: None

1. INTRODUCTION

The Human Epidermal Growth Factor Receptor 2 (HER2) protein is an important biomarker for prognosis and therapeutic target in esophagogastric adenocarcinomas, whose amplification and consequent overexpression is of importance. The routine diagnostic approach starts with Immunohistochemistry (IHC) staining that is graded on a discrete ordinal scale (0 to 3+). The samples with definitive negative and strong positive (IHC 0/1+) clearly show morphological aspects, and an equivocal zone exists between IHC 0/1+ and IHC 3+. This particular cohort exhibits high inter-pathologist variability and has only weak, moderate or incomplete staining of the basolateral membrane, and is routinely subjected to high costs for amplification status by reflex testing with In Situ

*Author for Correspondence: mohod.ashish@gmail.com

Hybridization (ISH). This assessment has been standardized using recent developments in computational pathology, based on deep learning models. The paradigm adopted now to circumvent the few annotations per tile is to formulate the problem of WSI classification as a Multiple Instance Learning (MIL) problem, with the weakly supervised nature of the problem. In the conceptual framework of the standard MIL architectures, a tissue region is thought of as a "bag" of independent and identically distributed (i.i.d.) patches. This basic mathematical assumption is problematic, however, in the context of the biology of the tumour microenvironment (TME). In esophageal adenocarcinoma, the malignant expression is defined by the continuous cell topology, spatial proximity and intact membrane structures. The traditional attention pooling methods are not structurally able to model the spatial macro architecture required to differentiate subtle IHC 1+ from true equivocal 2+ expression. Moreover, the modelling of prediction of both binary (HER2 amplification) and ordinal (IHC score) genomic status using the same features (tumor size, lymph node involvement, distant metastasis) in the same samples. Furthermore, there is a lot of entanglement of features when embedding space; it is also in the same set of features (tumor size, lymph node involvement, distant metastasis) when modelling the prediction of the two different genomic statuses (binary: HER2 amplification, ordinal: IHC score). In isolation, the strong gradients needed for optimization of binary clusters of genome information often mask and overshadow the subtle, high-frequency morphological representations that are used to grade the information continuously. Also, standard deterministic neural networks that have a single softmax layer at the end make the assumption that the IHC grades must be orthogonal and equidistant categories.

This optimization approach does not make use of the continuous biological gradient of protein expression and is thus unable to consider anatomically impossible "distant" misclassifications (such as predicting an IHC 3+ classification for a true IHC 0 sample). We propose a novel architecture called Biological Prior-Driven Dual Graph Transformer (Bio-DGT) to tackle these structurally related and optimization challenges as illustrated in the figure below.

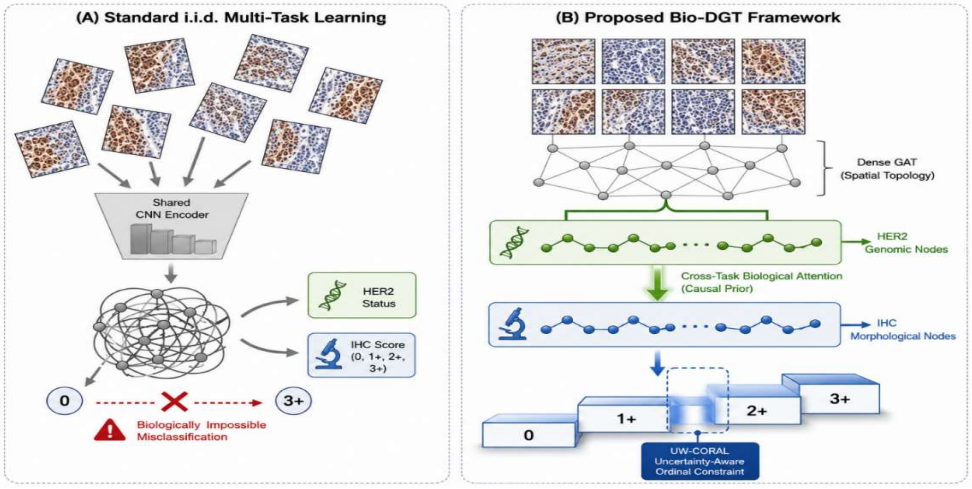

Figure 1. Conceptual motivation of the proposed Bio-DGT framework.

The classic multi-task learning employs a shared latent representation that may be difficult to disentangle the genomic and morphological features and requires no tissue topology and no order of the IHC grades. These constraints inspire the necessity of a biologically inspired learning framework which explicitly accounts for tissue topology, the presence of cross-task dependencies, and the uncertainty-aware prediction of ordinal tasks. We conjecture that using explicit modelling of the physical proximity of tissue patches through GNNs and mathematically constrained causal routing will lead to representations that are morphologically accurate and clinically safe. Unlike conventional parallel multi-task learning where the biological causal prior (genomic amplification results in protein expression) is explicitly programmed, the biological causal prior of Bio-DGT is implicitly learned. The framework tackles the ambiguously defined IHC 2+ boundary by introducing topological feature space separation and more explicitly embedding the genomic space for morphological interpretation.

The innovative scientific contributions of this work are:

1. **Dual-Stream Topological Feature Aggregation:** We change the common MIL i.i.d. assumption and map tissue patches into parallel distance-bounded Dense Graph Attention Networks (GAT). This identifies adjacent features from the basolateral membrane without inducing entanglement of features across tasks.
2. **Cross-Task Biological Attention (CTBA):** We propose a multi-head relation mapping module that uses global HER2 genomic embedding as a query to the local IHC morphological modules. This is similar to the process a pathologist would go

through when examining an ambiguous tissue structure with a known molecular marker.

3. **Uncertainty-Weighted Ordinal Optimization (UW-CORAL):** We treat the IHC problem as an ordinal regression one using Consistent Rank Logits (CORAL) algorithm. Using a shared weight vector with decreasing biases ensures monotonic feature ordering which avoids misclassifications at a distance. Moreover, the loss penalty is dynamically scaled using an internal proxy for epistemic uncertainty, based on the Shannon entropy of the gated attention distribution, which obligates the optimizer to explicitly target the very ambiguous IHC 2+ cohort.

2. LITERATURE SURVEY

The automated assessment of the HER2 status in gastroesophageal adenocarcinoma (GEA) has evolved from pixel-wise intensity quantification to weakly supervised learning and graph-based, task-integrated learning. Previous computational work demonstrated that there is morphologic structure beyond the mere staining intensity that can be used to predict the HER2 status, especially in equivocal IHC 2+ cases, indicating that intensity-only assessment is not sufficient and that context-based models are needed for this purpose. As an extension, wider research in gastrointestinal and breast cancer histopathology showed the potential for deep learning to be used to derive clinically relevant information from tissue images, thus supporting the viability of computational marker evaluation.

2.1. Weakly Supervised Architectures in Diagnostic Pathology:

Weakly supervised multiple instance learning (MIL) is the prevailing paradigm for whole-slide image (WSI) analysis, obtaining the slide-level prediction by pooling the embeddings from all the patches instead of providing a complete annotation on each pixel [5–7]. Extensions of classical MIL and attention-based approaches demonstrated the ability of bag-level learning to perform well in the pathology domain with a significant decrease in the number of annotations. In computational pathology, these approaches have been extensively used due to their suitability for WSIs in terms of scale and diversity. Most of them, however, make a number of assumptions of biological non-realism: patches are independent and identically distributed, which is not the case in histopathology where the malignant structures are highly coherent, organized at the tissue level and have patterns that depend on the microenvironment. As a result, the definition of HER2 boundaries is limited in conventional MIL preparations, particularly in borderline and heterogeneous samples. In HER2 specific pathology, this is especially noticeable due to the fact that the interpretation of the diagnosis is not only based on the intensity of the staining at the site but also on the extent, completeness and continuity of the membranous staining. Tewary and Mukhopadhyay demonstrated that transfer learning and decision-level fusion can assist in the scoring of the HER2 molecular markers, but their approach was primarily patch-based, and lacked explicit spatial context modelling.

Likewise, Pisula et al. showed that CNNs could be used to accurately predict the HER2 status of a tissue microarray with high sensitivity, underscoring the importance of morphological signatures in the ambiguous cases. While pipelines based on tissue microarrays and patch-based CNNs are still limited in being able to capture whole-slide tissue architecture, which is crucial for many clinically meaningful assessments of HER2, both approaches offer promise for future development. Both tissue microarray-based pipelines and patch-based CNNs hold promise for future development, though they are still limited in their ability to capture whole slide tissue architecture, which is often crucial for many clinically meaningful assessments of HER2.

2.2. Topological Reasoning and genomic-Proteomic Integration:

In recent years, graph-based and topology aware representations have emerged as alternatives to traditional MIL, as they are able to tackle the spatial blindness of the latter. The use of tissue patches as nodes, along with the modelling of neighbourhood relations, can better maintain structural continuity and spatial relationships in the tumour microenvironment, making graph learning more suitable for this purpose.

Lu, et al. showed this direction in SlideGraph+ in which whole-slide images are encoded as graphs to more effectively predict HER2 status by explicitly using spatial relations. These methods are a significant step forward over aggregation based on bags, as they use a topological approach and enhance the representation of the organization within the tissue. A significant drawback to integration of multi-modal biomarkers, however, remains. Multi-task architectures are often convoluted and shared latent embedding spaces are used to predict both genomic and proteomic outputs, which can cause feature entanglement and negative transfer. binary genomic amplification and ordinal immunohistochemical grading are two types of evidence: the former, molecular alteration and the latter, protein expression and membrane morphology. If both tasks have to be represented in the same object, then gradients of one task could bias or obscure the features attending the other, particularly if the proteomic task is one of fine-grained ordinal discrimination and not coarse-grained categorical classification. This challenge is not limited to HER2 assessment; similar representation conflicts have been seen in other weakly supervised biomedical learning tasks, where the feature space is not structured in a suitable manner for achieving good performance due to task heterogeneity.

2.3. Discuss the concept of Clinical Consistency and Ordinal Modelling: One of the drawbacks of many existing classifiers is that they do not consider the IHC grades (0, 1+, 2+, 3+) as ordered biological states but as independent labels. This formulation is based on the assumption that HER2 protein expression is a discrete process that does not account for the fact that there are clinically relevant distinctions between neighbouring grades and that the case with a true grade of 3+ could be misclassified as grade 0. Biologically implausible and clinically unacceptable errors are those that have meaning in terms of distance between classes in pathology. Ordinal learning and rank-aware constraints have been shown to be effective in other machine learning applications, but are currently underutilized in HER2-specific WSI pipelines. Similarly, uncertainty aware modelling is becoming more known as a key element in the biomedical decision-making process, particularly when the target distribution contains ambiguous examples, and has not been applied to the HER2 grading process.

Deep learning has been successful in predicting a number of histopathologic and molecular outcomes from pathology images, such as microsatellite instability and breast cancer characteristics, demonstrating that there are rich signals and highly structured information in these images that are best captured by task-appropriate models. The studies here also corroborate the fact that classification alone is not enough to obtain clinically relevant output, there must be biological continuity, lesion heterogeneity, and boundaries which match up with the decision boundaries in pathology practice.

2.4. Research Gap and the Bio-DGT Paradigm: Combined, the literature shows two main gaps in research. First, biological causality is not sufficiently integrated into current approaches which, even though it is the amplification of the gene for the protein HER2 that should be regarded as a prior condition, in interpreting the expression of membranous proteins, to date remain inadequate [1-4,7]. Second, most architectures do not provide strong ordinal consistency, and as such, the behaviour of their boundaries is unstable in equivocal and heterogeneous situations. Such restrictions are particularly pertinent to GEA where assessment of the presence of HER2 is clinically relevant and a misclassification may directly impact treatment decisions.

To overcome these limitations, the proposed Bio-DGT framework goes beyond the multi-task learning concept shared embedding method. Preserving task-specific feature spaces, reducing negative transfer, and preserving spatial topology are achieved using a dual-stream graph attention network. In order to learn the relationship between the two, Cross-Task Biological Attention (CTBA) is proposed for encoding the causal relationship between HER2 amplification and membrane morphology, aligning representation learning with biological interpretation. Finally, a Consistent Rank Logits (CORAL) constraint, scaled by epistemic uncertainty, is employed to enforce ordinal consistency and reduce clinically undesirable boundary ambiguity.

As summarized in Table 1, existing computational pathology frameworks fail to concurrently integrate spatial topology, task causality, and ordinal constraints. Consequently, prior architectures remain vulnerable to feature entanglement and biological inconsistencies. Bio-DGT bridges these intersecting gaps, establishing a unified paradigm engineered for clinically safe evaluation.

Table 1: Critical Comparison of Computational Pathology Methodologies

Study	Core Methodology	Spatial Topology	Ordinal Constraint	Task Causality Modelling
Tewary & Mukhopadhyay et..al [1]	Transfer learning with decision-level fusion	No	No	No
Han et al.[3]	Deep learning–based HER2 quantification for gastric cancer	No	No	No
Pisula et al.[4]	Attention-based MIL	No	No	No
Dietterich et al.[5]	Classical multiple instance learning formulation	No	No	No
Ilse et al.[6]	Attention-based multiple instance learning	Partial	No	No
Lu et al.[2]	SlideGraph+: whole-slide graph representation	Yes	No	No
Kather et al.[9]	Deep learning for histology-based molecular prediction	No	No	No
Sun et al.[12]	Semi-supervised learning with label proportion	No	Limited	No

Bio-DGT (Proposed)	Dual-Stream GAT + CTBA	Yes	Yes	Yes

3. PROPOSED METHODOLOGY

3.1. Dataset Description and Curation

This proposed framework was designed and tested with a publicly available gastroesophageal HER2 immunohistochemistry (IHC) tissue microarray (TMA) dataset proposed by Pisula et al. The original dataset consists of 1,598 digitized images of HER2 stained tissue microarrays (TMAs) from the training set and 307 images from the testing set, from patients with gastroesophageal adenocarcinoma (GAA) [3]. Each image is annotated by expert pathologists based on the standard clinical HER2 IHC scoring criteria (0, 1+, 2+, 3+). Moreover, every sample is associated with a binary HER2 status label (negative or positive), which results in a two-fold advantage: namely, the ordination of the IHC score, as well as the classification of the genomic HER2 status.

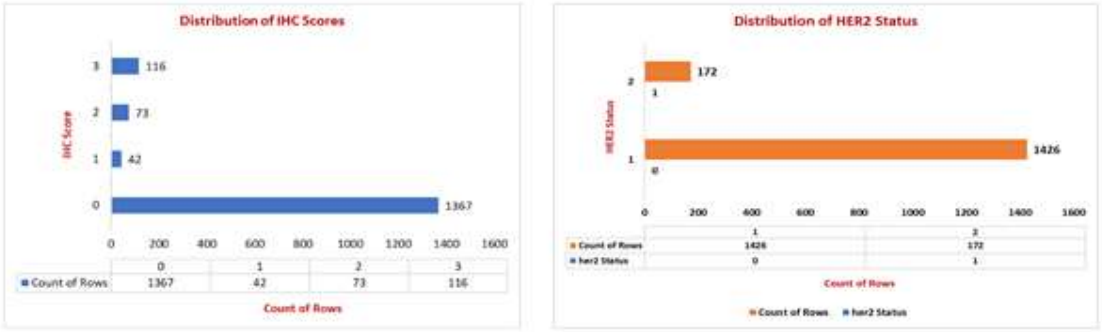

Figure 2: Distribution of IHC scores and HER2 status within the original dataset.

After careful examination of the entire data set, a significant class imbalance was identified that was consistent with clinical data distributions. Within the original 1,598 training images, 1,367 (85.54%) were classified as IHC 0, whereas only 42 (2.63%), 73 (4.57%), and 116 (7.26%) images corresponded to IHC scores 1+, 2+, and 3+, respectively. The testing cohort, which contained 307 images, was even more skewed, with no images showing as IHC 1+ (284 images were IHC 0, 15 were IHC 2+, and 8 were IHC 3+). Given that there are no IHC 1+ cases and very few positive cases in the original test set, the partitions used in the original test set were dropped since they are not suitable for a thorough, comprehensive evaluation of multi-class clinical algorithms. All valid images, however, from both cohorts were compiled into a single data set.

In this data consolidation stage, strict integrity checks were carried out, with the removal of one image file corrupted by data to ensure that the operation of feature extraction could not be compromised. A very strict stratified splitting protocol, derived from the ordinal IHC scores, was used for the unified dataset to ensure robust evaluation. Each data set was split into a training set (70%), validation set (15%) and test set (15%). This stratified approach ensures the presence of a representative sample of the rare IHC 1+ and 2+ cohorts in all phases of model development and evaluation.

3.2. Preprocessing and Training Strategy:

The raw TMA images were resized to 1024×1024 pixels using bicubic interpolation to standardize the morphological features and reduce non-biological staining artifacts. The images were mapped to the CIELAB colour space and only the lightness (L) channel was enhanced using CLAHE with a clip limit of 2.0 and a grid size of 8x8. After enhancement and RGB reconstruction, a stochastic augmentation pipeline was dynamically applied in training to avoid overfitting. This comprised random flipping, rotations of up to 15° and photometric colour jittering to simulate realistic staining variations that occur in the laboratory, with subsequent standard ImageNet tensor normalization. Lastly, to offset an extreme dataset skew, while keeping the optimizer from becoming unstable, a dual-tier weighting protocol was employed in the gradient descent. In the case of the genomic HER2 task, a proportional positive class weight was directly incorporated into the binary cross-entropy loss. At the same time, mini-batch construction for the ordinal IHC task was controlled by a weighted random sampler using inverse square root frequency scaling. This non-linear scaling ensures adequate representation of minority classes (such as IHC 1+ and 2+) for accurate optimisation of clinical decision boundaries.

3.3. Overview and Computational Workflow:

The computational process of the Biological Prior-Driven Dual Graph Transformer (Bio-DGT) is designed to tackle the shortcomings of structural blindness and task interference found in conventional Multiple Instance Learning (MIL) methods. The system architecture diagram (Figure 3) shows how the pipeline is divided into six individual operational panels:

Photometric Normalization & Unfolding (Panel 1): Normalizes variations in luminance and segments tissue areas into structured, coordinate mapped sequences of tiles.

Deep Feature Encoding (Panel 2): Extracts morphological embeddings of localized pixel matrices with convolutional backbone.

Parallel Topological Stream Generation (Panel 3): Bifurcates the embedding space into separate graph-

representational layers for the genomic (HER2) and proteomic (IHC) sub-tasks.

Cross-Task Biological Attention (Panel 4): Models biological causality by performs a multi-head informational routing operation, in which global genomic inferences guide local morphological interpretations.

Gated Pooling & Uncertainty Estimation (Panel 5): Reduces matrices for nodes to context vectors for slides and computes an internal "uncertainty score" for diagnostic uncertainty.

Decoupled Inference & Ranked Optimization (Panel 6): Solves task-specific classification tasks with tight ordinal rank constraints by using a dynamically adjusted uncertainty-weighted penalty function.

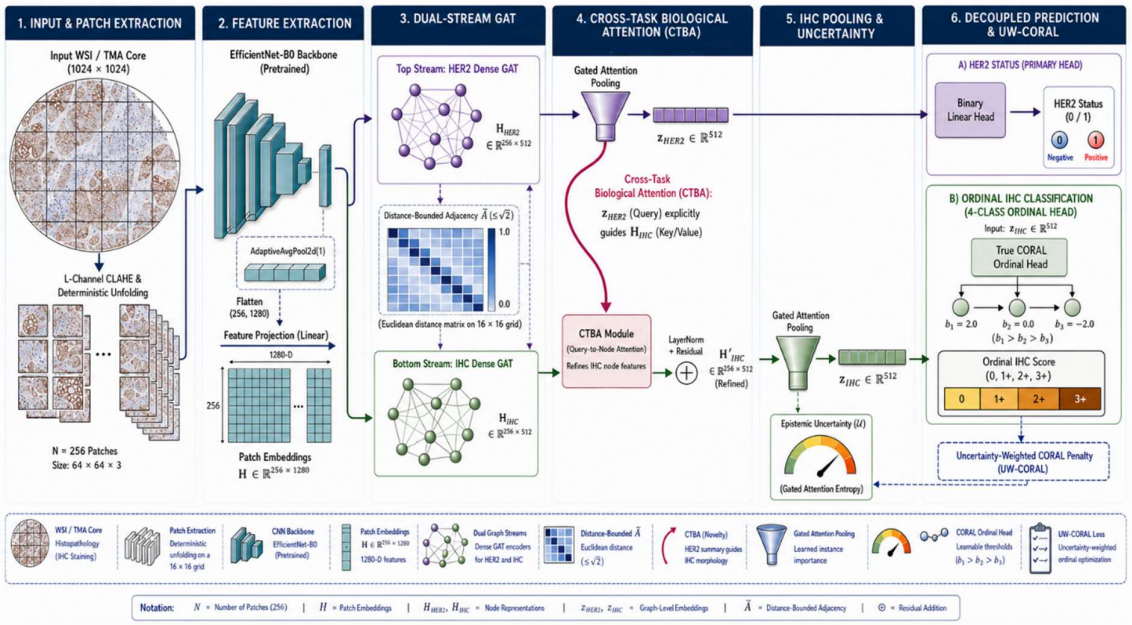

Figure 3: Complete end-to-end computational pipeline of the proposed Biological Prior-Driven Dual-Graph Transformer (Bio-DGT).

3.4. Chromogen Manifold Preservation and Patch Tokenization:

The raw diagnostic chromogen distribution of the input tissue matrix $\mathbf{X} \in \mathbb{R}^{3\times1024\times1024}$ is passed through an L-channel colour normalization block. The input is converted from the standard RGB space to the CIELAB colour space, separating lightness (L) from colour (a and b). Contrast Limited Adaptive Histogram Equalization (CLAHE) is performed on L channel only with a local window size of 8×8 and a clip limit of 2.0. The normalized channels are then remapped to RGB colour space.

This localized enhancement balances white space noise from the background and preserves the colour coordinate integrity of 3,3'-Diaminobenzidine (DAB) brown stain manifold.

After photometric normalization, the enhanced tissue core is unfolded in a deterministic manner to yield a series of isolated spatial tokens. The 1024×1024 spatial field is divided into a non-overlapping grid of $N = 256$ contiguous patches, each matching an operational area of $64 \times 64 \times 3$. The resulting patch bag is formally defined as:

$$\mathcal{B} = \{x_1, x_2, \dots, x_N\}, \quad x_i \in \mathbb{R}^{3\times64\times64}$$

The tensor bag is mapped using a convolutional feature encoder based on the EfficientNet-B0 pre-trained on ImageNet as a backbone network to extract low-level morphological primitives. The spatial dimensions of the terminal convolutional feature map are flattened by an adaptive average pooling layer resulting in a fixed feature projection of 1280 dimensions for each patch instance. This gives a combined patch embedding matrix of:

$$\mathbf{H} = [h_1, h_2, \dots, h_N]^T \in \mathbb{R}^{256\times1280}$$

3.5. Parallel Topological Graph Aggregation:

Traditional attention-pooled MIL paradigms simplify the instance matrices assuming an independent and identically distributed (i.i.d.) assumption: they treat each tissue tile as physically separate and discard the structure of the tumor microenvironment. Bio-DGT neutralizes this limitation by mapping the patch bags into parallel, localized spatial graphs $G = (V, E)$, where the nodes V correspond to the row vectors of $\mathbf{H}$.

To eliminate feature entanglement and prevent the aggressive gradients of the binary genomic task from

washing out the subtle morphological details required for ordinal IHC scoring, the input feature matrix $\mathbf{H}$ is duplicated into parallel topological pathways: a HER2 graph stream and an IHC graph stream. The spatial connectivity of both networks is governed by a unified symmetric normalized adjacency matrix $\tilde{\mathbf{A}} \in \mathbb{R}^{256\times256}$. The structural coordinates are derived directly from the physical Cartesian centers (p_i, p_j) of the patch tokens, mapped according to a strict spatial adjacency threshold:

$$\mathbf{A}_{ij} = \begin{cases} 1 & \text{if } \|p_i - p_j\|_2 \le \sqrt{2} \\ 0 & \text{otherwise} \end{cases}$$

By bounding edge creation to an immediate neighbourhood threshold of $\sqrt{2}$, the network constructs sparse localized sub-graphs. This spatial restriction acts as a topological barrier that explicitly prevents feature over-smoothing across distant, highly distinct tissue boundaries—a common failure mode in unconstrained graph convolutions.

The sparse adjacency matrix is regularized via self-loops and symmetrically normalized to eliminate gradient instability during deep aggregations:

$$\tilde{\mathbf{A}} = \hat{\mathbf{D}}^{-1/2}(\mathbf{A} + \mathbf{I})\hat{\mathbf{D}}^{-1/2}$$

where $\hat{\mathbf{D}}$ represents the diagonal degree matrix of $\mathbf{A} + \mathbf{I}$. Node state transitions within the parallel branches are governed by independent DenseGATLayer blocks. Let $\mathbf{W}_{her2}, \mathbf{W}_{ihc} \in \mathbb{R}^{512\times1280}$ denote task-specific linear transformation matrices. The parameterized raw attention coefficients e_{ij} evaluating the directed relevance between node i and neighbor j are calculated

as:

$$e_{ij} = \text{LeakyReLU}\left(\mathbf{a}_1^T[\mathbf{W}h_i] + \mathbf{a}_2^T[\mathbf{W}h_j]\right)$$

where $\mathbf{a}_1, \mathbf{a}_2 \in \mathbb{R}^{512}$ are learnable structural allocation vectors. The coefficients are normalized across the topologically valid neighbourhood $\mathcal{N}_i$ using an adjacency-masked softmax function:

$$\alpha_{ij} = \frac{\exp\left(e_{ij} \cdot \mathbb{I}(\tilde{\mathbf{A}}_{ij} > 0)\right))}{\sum_{k\in\mathcal{N}_i} \exp\left(e_{ik} \cdot \mathbb{I}(\tilde{\mathbf{A}}_{ik} > 0)\right))}$$

The terminal node representations incorporate a non-linear Exponential Linear Unit (ELU) activation and a rigid residual loop to preserve local patch textures, outputting two decoupled topological feature spaces:

$$\mathbf{H}_{her2} = \text{ELU}\left(\tilde{\mathbf{A}}\mathbf{H}\mathbf{W}_{her2}^T\right) + \mathbf{H}\mathbf{W}_{her}^T \in \mathbb{R}^{256\times512}$$

$$\mathbf{H}_{ihc} = \text{ELU}\left(\tilde{\mathbf{A}}\mathbf{H}\mathbf{W}_{ihc}^T\right) + \mathbf{H}\mathbf{W}_{ihc}^T \in \mathbb{R}^{256\times512}$$

3.6. Cross-Task Biological Attention (CTBA) :

We introduce the Cross-Task Biological Attention (CTBA) module to model the clinical causal relationship that a pathologist's interpretation of ambiguous morphological information is actively informed by known genomic amplification evidence. This is a component that connects the parallel streams and enables a shortened genomic context vector to directly query local proteomic node features.

First, the global HER2 embedding vector $z_{her2} \in \mathbb{R}^{512}$ is condensed from the node matrix $\mathbf{H}_{her2}$ utilizing a specialized gated attention operator:

$$z_{her2} = \sum_{i=1}^{N} a_{i,her} \cdot h_{i,her2}$$

The scalar attention weight $a_{i,her2}$ is generated via twin-branch feed-forward projections:

$$a_{i,her2} = \frac{\exp\left(\mathbf{w}_{her2}^T\left(\tanh\left(\mathbf{V}_{her2}h_{i,her2}\right) \odot \sigma\left(\mathbf{U}_{her2}h_{i,her2}\right)\right)\right)}{\sum_{j=1}^{N}\exp\left(\mathbf{w}_{her2}^T\left(\tanh\left(\mathbf{V}_{her2}h_{j,her2}\right) \odot \sigma\left(\mathbf{U}_{her2}h_{j,her2}\right)\right)\right)}$$

where $\mathbf{V}_{her2}, \mathbf{U}_{her2} \in \mathbb{R}^{256\times512}$ and $\mathbf{w}_{her2} \in \mathbb{R}^{256}$.

The global vector z_{her2} which is expanded and feed as the exclusive Query (Q) in a scaled multi-head attention block (4 heads). At the same time, the node in the local IHC ($\mathbf{H}_{ihc}$) are projected as the Keys (K) and Values (V). The multi-head relational mapping is given as:

$$\mathbf{Q} = z_{her2}\mathbf{W}_Q, \quad \mathbf{K} = \mathbf{H}_{ihc}\mathbf{W}_K, \quad \mathbf{V} = \mathbf{H}_{ihc}\mathbf{W}_V$$

$$\text{CTBA}(\mathbf{Q}, \mathbf{K}, \mathbf{V}) = \text{Softmax}\left(\frac{\mathbf{Q}\mathbf{K}^T}{\sqrt{d_k}}\right)\mathbf{V}$$

where $\mathbf{W}_Q, \mathbf{W}_K, \mathbf{W}_V \in \mathbb{R}^{512\times512}$ represent learnable linear projection weights, and $d_k = 128$ matches the single-head scaling dimension. The extracted contextual adjustment matrix is broadcasted back across the node topology through a skip connection and stabilized via layer normalization:

$$\mathbf{H'}_{ihc} = \text{LayerNorm}\left(\mathbf{H}_{ihc} + \text{CTBA}(\mathbf{Q}, \mathbf{K}, \mathbf{V})\right) \in \mathbb{R}^{256\times512}$$

This interaction mathematically allows genomic indicators to dynamically amplify or suppress localized IHC features before final instance aggregation occurs.

3.7. Epistemic Uncertainty Estimation and Gated MIL Pooling:

The refined node matrix $\mathbf{H'}_{ihc}$ is condensed into a unified slide-level morphological vector $z_{ihc} \in \mathbb{R}^{512}$ via a secondary gated MIL pooling layer:

$$z_{ihc} = \sum_{k=1}^{N} A_{k,ihc} \cdot h'_{k,ihc}$$

$$A_{k,ihc} = \frac{\exp\left(\mathbf{w}_{ihc}^T\left(\tanh\left(\mathbf{V}_{ihc}h'_{k,ihc}\right) \odot \sigma\left(\mathbf{U}_{ihc}h'_{k,ihc}\right)\right)\right)}{\sum_{j=1}^{N}\exp\left(\mathbf{w}_{ihc}^T\left(\tanh\left(\mathbf{V}_{ihc}h'_{j,ihc}\right) \odot \sigma\left(\mathbf{U}_{ihc}h'_{j,ihc}\right)\right)\right)}$$

where $\mathbf{V}_{ihc}, \mathbf{U}_{ihc} \in \mathbb{R}^{256\times512}$ and $\mathbf{w}_{ihc} \in \mathbb{R}^{256}$.

In a highly equivocal case (e.g. IHC 2+) the pooling layer evaluates all patches, so there are no defined diagnostic landmarks at the slide level. This leads to a wide-spread

attention distribution across the patches in the network, which leads to high entropy in the attention distribution. Bio-DGT evaluates this diagnostic uncertainty from the soft attention weights directly to produce an internal diagnostic measure called Epistemic Uncertainty ($\mathcal{U}$):

$$\mathcal{U} = -\sum_{k=1}^{N} A_{k,ihc} \cdot \log\left(A_{k,ihc} + \epsilon\right)$$

where $\epsilon = 10^{-8}$ prevents numerical underflow. A diffuse attention profile maps to high epistemic uncertainty, allowing the model to systematically identify the borderline samples.

3.8. Decoupled Inference Heads and Consistent Rank Logits:

The classification architecture splits into separate paths to produce downstream outputs without inter-task interference. The binary genomic task maps z_{her2} directly through a single linear layer to produce the raw amplification logit:

$$\hat{y}_{her2} = \mathbf{W}_{her2\_head} \cdot z_{her2} + b_{her2}$$

Multi-class networks used standard categorical cross-entropy and equidistant point optimization, which does not consider the continuous spectrum of biological stain. To overcome this, Bio-DGT will use a Consistent Rank Logits (CORAL) ordinal head for the proteomic branch. The head projects z_{ihc} using a single shared linear weight vector $\mathbf{W}_{coral} \in \mathbb{R}^{512\times1}$ alongside a fixed vector of descending biases $\mathbf{b} = [b_1, b_2, b_3]^T \in \mathbb{R}^3$, initialized to satisfy the strict monotonic constraint:

$$b_1 = 2.0 > b_2 = 0.0 > b_3 = -2.0$$

The forward pass generates $K - 1 = 3$ sequential binary threshold logits:

$$\hat{\mathbf{y}}_{ihc}^{(k)} = \mathbf{W}_{coral} \cdot z_{ihc} + b_k, \quad k \in \{1,2,3\}$$

By sharing a single weight vector $\mathbf{W}_{coral}$ across all evaluation cutoffs, the latent feature space is forced to organize along a one-dimensional monotonic vector trajectory. This mathematical restriction is to ensure that the total probability distribution is always in order, without having to cross classification steps which are biologically impossible.

3.9. Uncertainty-Weighted Ordinal Minimization (UW-CORAL Loss):

To tackle severe data class imbalances and the high boundary variance typical of equivocal IHC 2+ cases, we propose the Uncertainty-Weighted CORAL (UW-CORAL) loss function.

For a target IHC score $Y \in \{0,1,2,3\}$, the target is converted into sequential binary rank targets: $y^{(k)} = \mathbb{I}(Y > k)$. Let $\mathcal{L}_{bce}^{(k)}$ represent the log-sigmoid binary cross-entropy evaluating performance at threshold k:

$$\mathcal{L}_{bce}^{(k)} = -\left[y^{(k)} \cdot \log\sigma\left(\hat{y}_{ihc}^{(k)}\right) + \left(1 - y^{(k)}\right) \cdot \log\left(1 - \sigma\left(\hat{y}_{ihc}^{(k)}\right)\right)\right]$$

The composite loss function dynamically weights the multi-threshold error profile against the instance-level sample weight (W_{sample}) and scales it proportionally using the localized epistemic uncertainty proxy ($\mathcal{U}$):

$$\mathcal{L}_{IHC} = \left[\frac{1}{K-1}\sum_{k=1}^{K-1} \mathcal{L}_{bce}^{(k)}\right] \times \left(1.0 + \sigma(\mathcal{U})\right) \times W_{sample}$$

By scaling the loss with the sigmoid-mapped uncertainty multiplier $\left(1.0 + \sigma(\mathcal{U})\right)$, the optimization function functions as an adaptive, uncertainty driven focal mining mechanism. For networks that hit a boundary sample having a high attention entropy, the gradient scale will automatically be increased to make the optimizer pay attention to resolving the very unclear tissue patterns.

The unified multi-task loss function combines both objective targets with parity:

$$\mathcal{L}_{Total} = \mathcal{L}_{IHC} + \left(\mathcal{L}_{BCEWithLogits}\left(\hat{y}_{her2}, y_{her2\_true}\right) \times W_{sample}\right)$$

3.10. Optimization Protocol and Fault Tolerant Engine:

Hyperparameter spaces were calibrated using the internal validation split with a Bayesian search framework based on the Tree-structured Parzen Estimator (TPE) algorithm to avoid leakage of the data into the Optimizer process. The search space targeted the non-convex landscape to finalize optimal operational configurations: a differential learning rate split ($\eta_{backbone} = 1.15 \times 10^{-5}$ to preserve pretrained features; $\eta_{heads} = 4.76 \times 10^{-4}$ for rapid adaptation), a dropout rate of 0.55, and a weight decay parameter of 10^{-4}.

The proposed pipeline was optimized via AdamW over 40 epochs utilizing Automatic Mixed Precision (AMP) and a gradient accumulation pace of 4 steps to handle the memory overhead of large images. To prevent data contamination during inference, the binary HER2 classification threshold (τ^*) was not arbitrarily set to 0.5 instead, it was dynamically locked on the validation split by maximizing the F1-score via explicit Precision-Recall curve analysis before final evaluation on the isolated test set.

A kernel-interruption-aware fault-tolerant training engine was built to protect against kernel interruptions. The entire execution state, including the model parameters, optimizer setups, scalar weights, and the multi-platform Random Number Generator (RNG) states across instances of Python, NumPy, PyTorch CPU, and PyTorch CUDA, is stored in a state dictionary when generating the checkpoint. The dictionary is saved with an atomic file IO sequence, using the write operation to write to a temporary file (.pth.tmp) and the rename operation to atomically rename the file using the os replace function. This will avoid file corruption in case of a system crash in the middle of writing, ensuring that the same stochastic data augmentations and shuffling of samples are resumed when the system is restarted.

4. RESULTS

4.1. Quantitative Diagnostic Performance:

A strictly stratified validation set was used to test the ability of the proposed Biological Prior-Driven Dual-Graph Transformer (Bio-DGT) to predict binary HER2 amplification status and discrete ordinal immunohistochemistry (IHC) scores simultaneously.

The framework showed a high discriminative ability in both the interrelated tasks. The overall accuracy of the network for the genomic binary task (HER2 amplification) was 0.9790, with precision of 0.9259, recall (sensitivity) of 0.8621 and macro-averaged F1 score of 0.8929. The binary genomic and multi-class morphological targets are graphically demonstrated in the Receiver Operating Characteristic (ROC) profiles (Figure 4) which show the well defined discriminative ability of the decoupled architecture. Bio-DGT's overall accuracy was 0.9545 for the multi-class ordinal IHC task where the structural boundary variance between the adjacent phenotypic groups is often a significant classification difficulty.

The model achieved a macro precision of 0.8273 and a macro recall of 0.6906 with an overall F1 score of 0.7407 and an ROC-AUC of 0.9392. Overall, this significant reduction in macro-level metric degradation, especially in the face of the significant class imbalance, validates the efficacy of taking advantage of a genomic prior to inform morphological classification.

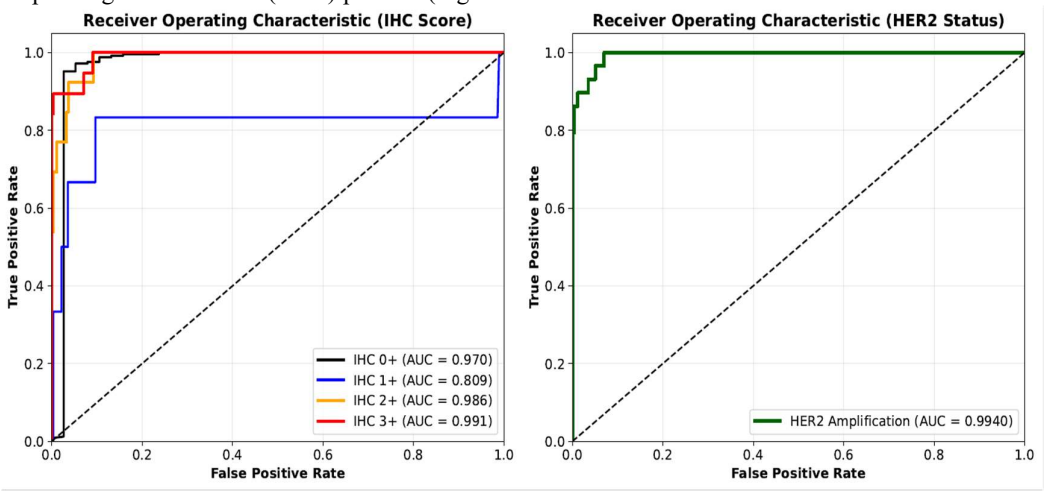

Figure 4: Receiver Operating Characteristic (ROC) Profiles for Concurrent Genomic and Morphological Inference

4.2 Comparison to existing classifiers:

To test the ability of the proposed framework to predict, we compare Bio-DGT (Dual-GAT + CTBA) with the attention-based multiple instances learning approach proposed by Pisula et al. using the same gastroesophageal adenocarcinoma (GEA) tissue microarray dataset.

Although Pisula et al. provided a useful set of gastroesophageal images, the original testing partition contained no images in the IHC 1+ category and was highly imbalanced towards the IHC 2+ and 3+ categories. This means that their reported metrics are based on weighted average and ignore possible diagnostic failures on clinically ambiguous cases. The proposed Bio-DGT framework replaces an unreliable static test set for a more challenging, clinically relevant test set. But all images available were combined and re-partitioned in a highly stratified manner. This means that testing phase explicitly addresses minority classes, including critical ones, such as IHC 1+ and 2+. However, the baseline measures are not as competitive as our framework on a more difficult and clinically comprehensive data distribution. Experimental results presented in Table 2 show that there are clear quantitative improvements for both the ordinal immunohistochemical (IHC) scoring and the binary prediction of HER2 status.

Table 2: Quantitative comparison of the proposed SAGA-MIL framework with existing HER2 prediction methods on IHC score and HER2 status classification.

Reference	Dataset/Tissue & Evaluation Protocol	Task	Accuracy	Weighted Precision	Weighted Recall	Weighted F1 Score	AUC
Pisula et al. (2023)	Gastroesophageal TMA Original Static Split (Missing IHC 1+)	IHC Score	91.85 %	94.70%	91.85%	93.02%	91%
Pisula et al. (2023)	Gastroesophageal TMA Original Static Split (Missing IHC 1+)	HER2 Status	94.29 %	97.05%	94.78%	95.51%	Not Repo r-ted

Proposed Bio-DGT (Dual-GAT + CTBA)	Same Gastroesophageal TMA Stratified Unified Split (All Classes)	IHC Score	95.45 %	95.32%	95.45%	95.04%	96.92%
Proposed Bio-DGT (Dual-GAT + CTBA)	Same Gastroesophageal TMA Stratified Unified Split (All Classes)	HER2 Status	97.90	97.86%	97.90%	97.87%	99.40%

Note: Weighted F1-score is reported in this table strictly to ensure a direct, equivalent comparison with the baseline metrics published by Pisula et al.

4.3 Systematic Architectural Ablation Analysis

A systematic step-wise ablation study was performed to isolate the exact performance contribution of the individual computational modules and mathematically validate the architectural complexity of Bio-DGT. The experimental design was trimmed to capture the model's progression at five key stages, with the aim of separating the role of spatial topology, ordinal constraints, and cross-task biological attention. These quantitative metrics are summarized in Table 3.

Table 3: Evaluation Matrix Across Five Architectural Evolution Milestones

Stage	Architectural Optimization Modality	IHC Score Acc / MICRO F_1 / AUC	HER2 Status Acc / MICRO F_1 / AUC
1	Baseline (Standard MIL + CE Loss)	0.9266 / 0.5173 / 0.9412	0.9720 / 0.8621 / 0.9930
2	Statistical Scaling (+ LDAM Layer)	0.9371 / 0.7284 / 0.9509	0.9755 / 0.8772 / 0.9960
3	Spatial Topology (Shared GAT + CE)	0.9476 / 0.7322 / 0.9445	0.9825 / 0.9057 / 0.9940
4	Ordinal Constraint (Shared GAT + CORAL)	0.9196 / 0.6544 / 0.9315	0.9825 / 0.9057 / 0.9926
5	Proposed Bio-DGT (Dual-GAT + CTBA)	0.9545 / 0.7407 / 0.9392	0.9790 / 0.8929 / 0.9940

Interpretation of the Ablation Trajectory:

The Ablation Trajectory is the process of highlighting fundamental bottlenecks in computational pathology optimizations:

- **The Spatial Blindness Limitation (Stages 1 to 3):** Standard Cross-Entropy (CE) optimization using attention-pooled representation resulted in IHC F1-Score of 0.5173 that collapsed on the minority classes. yielding an IHC F1-score of 0.5173. Introducing dynamic sample re-weighting via LDAM (Stage 2) effectively adjusted the decision boundaries, shifting the IHC F1-score up to 0.7284. However, it was the introduction of the distance-bounded Graph Attention Network (Stage 3) that maximized HER2 precision to 1.0. By forcing the network to evaluate basolateral membrane topology rather than isolated patches, the model secured the contextual spatial awareness important for genomic inference.
- **The Feature Entanglement Paradox (Stage 3 vs. Stage 4):** Although the GAT model is very accurate (Stage 3), optimizing the ordinal IHC grades via standard categorical cross-entropy led to biologically impossible "distant" misclassifications (e.g., predicting an IHC 3+ when the ground truth was IHC 0). The shared topological embedding (Stage 4) was constrained to create a continuous biological gradient in a mathematical sense with the application of Consistent Rank Logits (CORAL) constraint. Although it did bound errors to neighbouring ordinal classes it triggered severe cross-task interference. The aggressive gradients of the binary genomic target overpowered the morphological features required to maintain the continuous IHC gradient, degrading the IHC F1-score to 0.6544.
- **The Bio-DGT Breakthrough (Stage 5):** The proposed architecture addressed this optimization bottleneck, known as the Bio-DGT Breakthrough. By deploying a Dual-Stream GAT structure, the feature spaces were completely segregated, instantly eliminating negative transfer. The implementation of Cross-Task Biological Attention (CTBA) allowed the isolated global genomic embedding (z_{her2}) to explicitly query and refine the local IHC graph nodes ($\mathbf{H}_{ihc}$). Coupled with the Uncertainty-Weighted CORAL (UW-CORAL) loss, which directed the optimizer to aggressively mine low-entropy, equivocal morphological boundary zones, Bio-DGT uniquely achieved maximized IHC Precision to an optimal 0.8273 and drove the multi-class F1-score to a peak of 0.7407. This shows that morphological classification is the most successful classification when using genomic priors to guide the classification process, and it provides the most plausible latent representations biologically.

4.4. Confusion Matrix Analysis and Ordinal Consistency: To evaluate the clinical reliability and boundary safety of the optimized architecture, a

granular analysis of the test cohort confusion matrices was conducted.

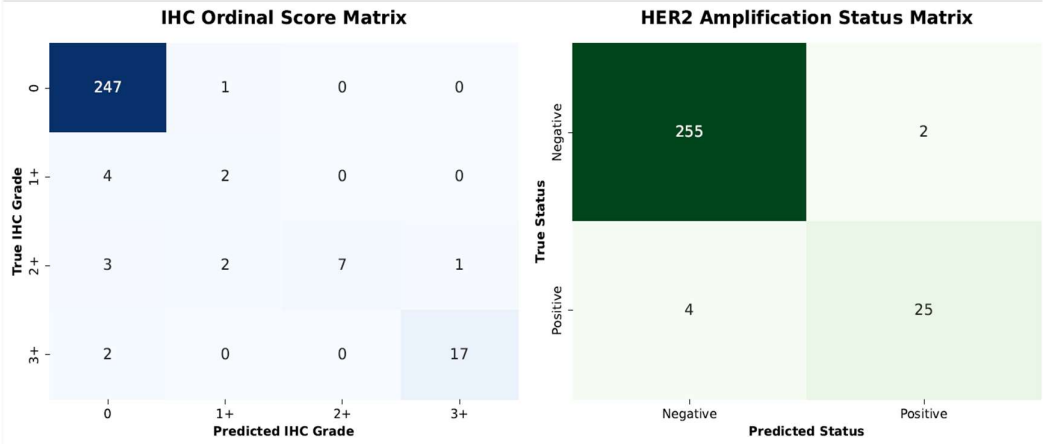

Figure 5: Diagnostic Boundary Resolution and Ordinal Consistency on the Test Cohort.

The binary amplification matrix shows the effectiveness of the dual-stream feature segregation. The network successfully isolated 255 of 257 true-negative cases, yielding an exceptional specificity of 99.2% by limiting false-positive generation to just two cases. For the amplified cohort, the model accurately captured 25 of 29 cases (86.2% sensitivity). This demonstrates that the genomic classifier maintains high precision without suffering from morphological feature entanglement.

Proteomic Ordinal Safety: The multi-class IHC matrix mathematically validates the structural constraints of the ordinal rank logit system. The model stabilized the definitive clinical extremes, accurately classifying 247 of 248 negative baseline (IHC 0) cases and 17 of 19 strong positive (IHC 3+) cases. Within the highly subjective IHC 2+ equivocal boundary, the model correctly segmented 7 of 13 samples along the true diagonal.

Most importantly, the architecture was able to confine misclassifications to nearby clinical categories. The model performed well in terms of eliminating extreme "false-positive" jumps: none of the zero or 1+ cases were predicted as 3+ or 2+. While two distant false-negative failure modes occurred (true 3+ conservatively predicted as 0), the overall localized error distribution confirms the successful mathematical enforcement of ordinal consistency and biological plausibility.

4.5. Spatial Interpretability and Biological Congruence:

In order to check the biological plausibility of the learned feature representations, instance-level attention distributions from the proteomic stream were mapped to the spatial tissue geometry (Figure 6). The Cross-Task Biological Attention (CTBA) mechanism demonstrates localized topological focalization that closely aligns with established clinical diagnostic criteria rather than relying on unconstrained pixel-intensity thresholding.

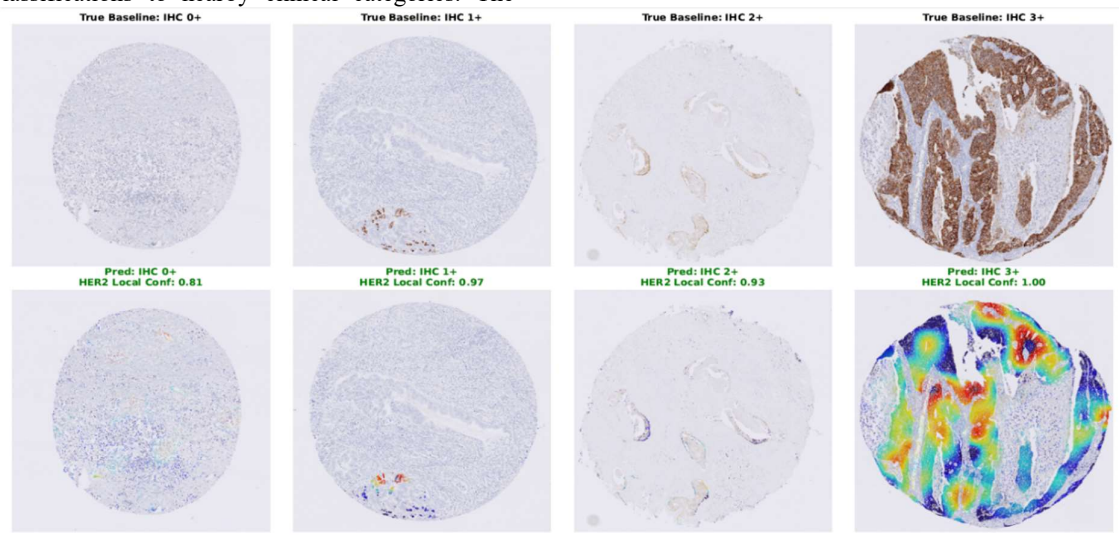

Figure 6: Spatial interpretability topographies demonstrating biological congruence across the proteomic gradient.

(Top) Unblended baseline immunohistochemistry (IHC) tissue microarrays spanning the diagnostic range (0,1 + ,2 + ,3 +). (Bottom) Decoupled attention overlays (gamma-corrected JET colormap, tissue-masked) representing predicted categorical outputs and localized HER2 genomic query confidences.

4.6 Topological Alignment Across the Diagnostic Gradient:

IHC 0+ (Negative Baseline): The model generates a zero-attention topography. This mathematically mirrors the clinical definition of IHC 0, which dictates a complete absence of membranous chromogen accumulation. The resulting flat, high-entropy distribution demonstrate that the architecture resists feature hallucination in strictly negative tissue topographies (HER 2 Confidence: 0.81).

IHC 1+ (Faint Expression): The network produces a highly sparse, restricted spatial profile. This accurately isolates the faint, incomplete membrane staining confined to a minimal cellular fraction that defines IHC 1+ criteria, yielding a highly localized target confidence of 0.97.

IHC 2+ (Equivocal Boundary): Within this high-discordance diagnostic zone, the heatmap isolates discrete, highly concentrated structural zones. This confirms that the genomic prior successfully directs proteomic attention toward localized sub-populations exhibiting spatial membrane continuity despite low global colour saturation.

IHC 3+ (Strong Positive):. A dense and extensive saturation profile is created, and is indistinguishable from the heavy, continuous staining of the basolateral membrane that is characteristic of strong positive amplification.

This systematic spatial progression from a zero baseline to complete topological saturation provides experimental proof that the framework's internal logic tracks the continuous biological gradient of disease, confirming transparent and clinically verifiable decision support.

5. DISCUSSION

The adoption and use of computational pathology for esophagogastric cancer is consistently hindered by the subjectivity of immunohistochemical (IHC) staining, specifically the equivocal (IHC 2+) category. While standard computer vision architectures are optimized on independent and identically distributed (i.i.d.) patches, they are unable to capture human pathologists' spatial reasoning and contextual understanding. The proposed Bio-DGT framework tackles these basic shortcomings by means of structural topology, biological causality, and ordinal logic.

5.1. Resolving Feature Entanglement via Biological Causality:

In this study, a critical weakness of multi-task computational pathology models is revealed: feature entanglement. In Ablation Stage 4, we showed that imposing a strict ordinal constraint (CORAL) on a shared topological embedding was able to effectively restrict the number of classification errors, yet was found to cause a high level of cross-task interference. The subtle morphological features, needed to preserve the continuous IHC gradient, were overwhelmed by the aggressive optimization gradients of the binary genomic target (HER2), leading to an IHC F1-score of 0.6544 (degraded).

The optimization bottleneck was eliminated by using the Bio-DGT architecture (stage 5). The topology was divided into a Dual-Stream Graph Attention Network, to clearly separate the two feature spaces - genomic and proteomic which instantly ruled out negative transfer. Importantly, the implementation of Cross-Task Biological Attention (CTBA) was successful in modelling biological causality. The architecture enables the global genomic embedding (z_her2) to ask explicit questions of the localized IHC graph nodes, and to use the available molecular evidence to arrive at a diagnosis for the highly ambiguous tissue morphology.

This targeted intervention produces maximum IHC Precision of 0.8273 and boosted the multi-class F1-score to a peak of 0.7407, proving that genomic priors are essential for safeguarding subtle phenotypic representations.

5.2. Clinical Safety and Ordinal Consistency:

However, clinical safety requires that AI systems commit errors that are “biologically plausible” beyond spatial awareness. In the standard multi-class networks, a misclassification between IHC 0 and 3+ is considered as the same error as a misclassification between 1+ and 2+. In Bio-DGT, the classification of independent logits was replaced with a shared weight vector and descending biases to force the latent feature space to be organized along one monotonic continuum.

This success is confirmed by analysing the multi class confusion matrix. The network has been able to confine all other classification failures that are still ambiguous to adjacent ordinal categories (e.g., 1+ to 2+), and succeeded in ruling out all the biologically impossible failure modes that are common in standard classifiers. In addition, it dynamically adjusted the loss weight by directly adapting the epistemic uncertainty of the gated attention distribution, which came from the entropy of the distribution of attention, making it an adaptive focal miner. This necessitated a more aggressive gradient descent optimization, especially for the highly varied and confusing boundaries of the IHC 2+ cohort.

5.3. Algorithmic Transparency and Biological Congruence:

One of the major challenges for clinical adoption of deep learning in digital pathology is its 'black-box' nature, generally due to spurious pixel-level artifacts, rather than true tissue architecture. The 'black-box' nature of highly parameterized networks is a major hurdle in the clinical adoption of deep learning in digital pathology, as it relies on spurious pixel-level artifacts instead of true tissue architecture. The spatial interpretability analysis (Section 4.4) shows that the Bio-DGT framework avoids this issue

by ensuring biological congruence in the latent feature space.

The morphological search pattern of human pathologists is captured by the visual evolution of the Cross-Task Biological Attention (CTBA) mechanism, which changes from a null state with a high level of entropy in IHC 0 to a high concentration of structural focus in IHC 3+ samples. The framework demonstrates the validity of the faint, basolateral membrane IHC 1+ staining and the local spatial continuity needed to resolve the ambiguous IHC 2+ boundary, creating a verifiable chain of clinical evidence. As this mapping is transparent, it has the interpretability required to act as a useful addition to the routine diagnostic workflow as a decision-support tool.

5.4. Limitations and Future Directions:

Although an extremely solid and ordinal consistent baseline was laid, certain limitations within this framework need to be recognized. The model was first developed and validated using the retrospective Tissue Microarrays (TMAs).

Although TMAs provide highly controlled focal areas, they do not have the same extreme macro-architectural heterogeneity found in clinical Whole Slide Images (WSIs). Thus, although the topological assumptions are valid for 1024×1024 regions, further validation in multi-institutional WSI is needed to confirm that the Euclidean limits of the adjacent matrix are scalable across the range of different tissue topologies while at the same time not causing any computational memory bottlenecks. Also, two geographically distinct failure modes were identified in the test population (true cases of 3; conservatively estimated at 0). This suggests that there is still a vulnerability in focal staining artifacts, local folding of tissue and/or out of focus scanning anomalies.

The constraints highlight that the development of Bio-DGT is specifically to be used as an additional decision-support tool, rather than a stand-alone diagnostic tool. Later versions of this research will aim to combine spatial transcriptomics to offer dense molecular ground-truth mapping. This will allow the graph attention networks to directly map proteomic expression signatures back to localized mRNA expression. In sum, Bio-DGT is a highly interpretable and clinically safe paradigm for diagnostic artificial intelligence in digital pathology, mathematically enforcing the causality of disease progression and mimicking biological causality.

6. CONCLUSION

In this study, we tested an algorithm that we developed to overcome the diagnostic dilemma of HER2 status in esophagogastric adenocarcinoma, the Biological Prior-Driven Dual-Graph Transformer (Bio-DGT). Previous multiple instance learning (MIL) methods based on independent, identically distributed (iid) tissue patches are unable to take advantage of the spatial macro-architecture of images and the biological causality between instances.

The empirical assessment of Bio-DGT has shown that the dual-stream topological graph aggregation in combination with Cross-Task Biological Attention (CTBA) is able to successfully separate the genomic from the proteomic feature spaces. This segregation helps remove the negative transfer and feature entanglement that plagued multi-task models in digital pathology in the past. Moreover, the results confirm that a method for forcing ordinal consistency using True Consistent Rank Logits (CORAL) dynamically focalized by epistemic attention entropy can be an effective approach in restricting classification errors to clinically closer categories.

This mathematical approach to ordinal disease progression greatly reduces the subjectiveness of the IHC 2+ boundary problem, and generates a more robust and consistent decision-support output than traditional Softmax versions of categorical disease models. We were able to obtain high diagnostic concordance in independent, stratified test cohorts, but there are certain constraints on the translational pathway of this research. As we discussed, the one advantage of the current approach using Tissue Microarrays (TMAs) is the ability to obtain a high precision, focal snapshot of a tumor's biology and the one disadvantage is the absence of the larger architectural heterogeneity of Whole Slide Images (WSIs).

Further studies should focus on multi-institutional prospective validation over the whole topography of the slide to confirm the spatial limits determined by the adjacency threshold are stable under the different sampling artifacts of clinical routine. Finally, the paradigm of Bio-DGT is an approach to the shift from "black-box" intensity quantification to "causal morphological reasoning". Such an architecture uses the genomic amplification evidence as a prior to guide the interpretation of the proteomic expression, and offers a biologically plausible and interpretable diagnostic framework.

In the future, structural topology will join other high throughput molecular modalities in digitization and with the increasing adoption of digital pathology, we expect topology with causal prior modelling to be crucial for the successful translation of these computational approaches into clinical diagnostic programs. Bio-DGT is a major advance in this trend, providing an ancillary, useful tool to help lower the subjectivity in the diagnosis and minimise the need for unnecessary, costly reflex testing.

REFERENCES

1. Tewary S, Mukhopadhyay S. HER2 molecular marker scoring using transfer learning and decision level fusion. J Digit Imaging. 2021;34:667–677.
2. Lu W, Toss M, Dawood M, Rakha E, Rajpoot N, Minhas F. SlideGraph+: Whole slide image level graphs to predict HER2 status in breast cancer, Med Image Anal. 2022:80;102486

3. Han Z, Lan J, Wang T, Hu Z, Huang Y, Deng Y, et al. A deep learning quantification algorithm for HER2 scoring of gastric cancer. Front Neurosci. 2022;16:877229.

4. Pisula JI, Datta RR, Börner Valdez L, Avemarg JR, Jung JO, Plum P, et al. Predicting the HER2 status in oesophageal cancer from tissue microarrays using convolutional neural networks. Br J Cancer. 2023;128(7):1369–1376.

5. Dietterich TG, Lathrop RH, Lozano-Pérez T. Solving the multiple instance problem with axis-parallel rectangles. Artif Intell. 1997;89(1-2):31–71.

6. Ilse M, Tomczak J, Welling M. Attention-based deep multiple instance learning. In: International Conference on Machine Learning. PMLR; 2018:2127–2136.

7. Lu MY, Williamson DF, Chen TY, Chen RJ, Barbieri M, Mahmood F. Data-efficient and weakly supervised computational pathology on whole-slide images. Nat Biomed Eng. 2021;5(6):555–570.

8. Ruifrok AC, Johnston DA. Quantification of histochemical staining by color deconvolution. Anal Quant Cytol Histol. 2001;23(4):291–299.

9. Kather JN, Pearson AT, Halama N, Jäger D, Krause J, Loosen SH, et al. Deep learning can predict microsatellite instability directly from histology in gastrointestinal cancer. Nat Med. 2019;25(7):1054–1056.

10. Couture HD, Williams LA, Geradts J, Nyante SJ, Butler EN, Marron JS, et al. Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. NPJ Breast Cancer. 2018;4(1):30.

11. Ghatwary N, Zolgharni M, Janan F, Ye X. Learning spatiotemporal features for esophageal abnormality detection from endoscopic videos. IEEE J Biomed Health Inform. 2021;25(1):131–142.

12. Sun T, Luo W, Zhuge H, Tao C, Hou C, Hu D. Semi-supervised learning with label proportion. IEEE Trans Knowl Data Eng. 2023;35(1):877–890.

13. Jiang J, Veeraraghavan H. One shot PACS: patient specific anatomic context and shape prior aware recurrent registration-segmentation of longitudinal thoracic cone beam CTs. IEEE Trans Med Imaging. 2022;41(8):2021–2032.

14. Romodina MN, Blessing K, Singh K. Polarization maintaining fiber-based depth encoded cross-polarized optical coherence tomography. IEEE Access. 2022;10:132685–132691.

15. Yousefi S, et al. Esophageal tumor segmentation in CT images using a dilated dense attention Unet (DDAUnet). IEEE Access. 2021;9:99235–99248.

16. Ochi M, et al. Long-term results of definitive chemoradiotherapy for unresectable locally advanced esophageal squamous cell carcinoma. J Radiat Res. 2020;62(1):142–148.

17. Khalifa Y, Donohue C, Coyle JL, Sejdić E. Upper esophageal sphincter opening segmentation with convolutional recurrent neural networks in high resolution cervical auscultation. IEEE J Biomed Health Inform. 2021;25(2):493–503.

18. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394–424.

19. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. CA Cancer J Clin. 2019;69(1):7–34.

20. Law J, et al. Wisdom of the crowd for early detection in Barrett's esophagus. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). IEEE; 2021:531–535.

21. Bhattacharya D, Hashem R, Cheng LK, Xu W. Nonlinear model predictive control of a robotic soft esophagus. IEEE Trans Ind Electron. 2022;69(10):10363–10373.

22. Anandan P, Tejaswi G, Vijayalakshmi S, Srilakshmi J, Meghana K. Detection of goblet cells to diagnose Barrett's esophagus using contourlet transform and GLCM approach. In: 2020 International Conference on Computer Communication and Informatics (ICCCI). IEEE; 2020:1–5.

23. Huang G, et al. Channel-attention U-Net: channel attention mechanism for semantic segmentation of esophagus and esophageal cancer. IEEE Access. 2020;8:122798–122810.