- Wojciechowska, M., Malacrino, S., Garcia Martin, N., Fehri, H., & Rittscher, J. (2021, September). **Early Detection of Liver Fibrosis Using Graph Convolutional Networks.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 217-226). Springer, Cham.
- Yin, C., Liu, S., Shao, R., & Yuen, P. C. (2021, September). **Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification.** In *International Conference on Medical Image Computing and Computer-Assisted Intervention* (pp. 153-162). Springer, Cham.

2021.11.10 Lufei GAO

Background

Early Detection of Liver Fibrosis Using Graph Convolutional Networks

- Assessing the level of fibrosis in the diagnosis of liver disease in general and **Non-alcoholic Steatohepatitis (NASH)** in particular.
- **The measurement of the Collagen Proportionate Area (CPA)** provides a very basic assessment of the collagen content per unit area, it has already been shown that **the assessment of morphometrical features** allows for a more accurate histological staging of fibrosis.
- Deep learning to identify collagen deposition patterns
  - Identifies the macro-patterns and characterizes the different fibrosis stages through **modelling the level of bridging**.
- GCNs apply the convolutions to **predefined graph nodes**.
  - represent **biological entities** as in the case of cell-graphs, or larger tissue regions – super pixels or groups of tiles of similar appearance.
- A novel **tissue representation** using **graph-based model**
  - **robustness of deep learning + prior knowledge of tissue architecture**
  - Resemble a human reader's understanding of the slide, leading to **more interpretable analysis** of fibrosis.

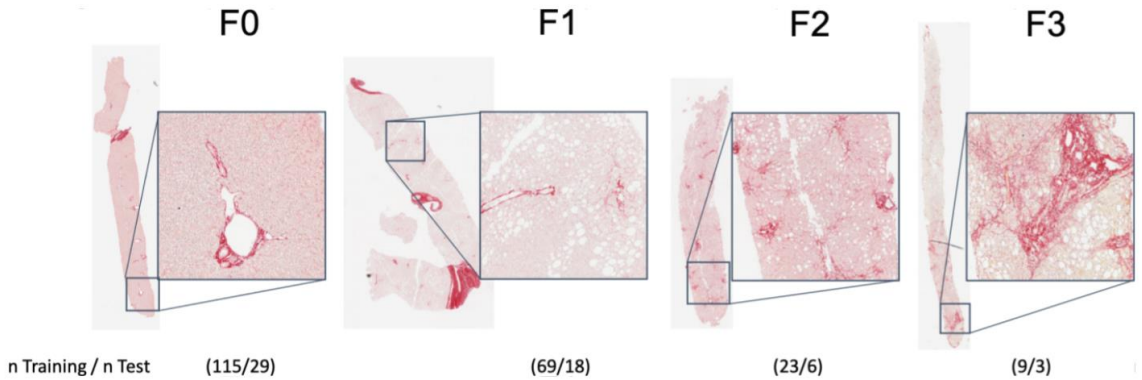
Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification.

- Nonalcoholic fatty liver disease (NAFLD)
- Liver biopsy images: quantitative analysis of histological patterns
  - **fibrosis**
  - 3 NAS-related components (**ballooning degeneration**, **lobular inflammation**, and **steatosis**)
  - The analysis of NAS-related components is more challenging and rarely studied
- Develop an automatic image analysis model to assist pathologists
  - A **selective attention regularization method** to clearly utilize clinically interpretable features for the analysis of NAS-related components.
    - **Selective attention regularization module (SAttenReg)**
    - Explicitly drives the model to focus on clinically interpretable features (e.g., nuclei and fat droplets) to improve the interpretability and reliability of the model

Datasets

Early Detection of Liver Fibrosis Using Graph Convolutional Networks

- provided by Perspectum Ltd: A set of **271 percutaneous liver biopsies** from patients with NAFLD
  - Representing varying stages of fibrosis
  - Graded by the METAVIR standard (F0-F4)
  - Divided into training and testing folds with a 80:20 ratio.
  - F4 is not presented in the dataset.



Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification.

- **Biopsy4Grading**
  - Public liver section dataset collected from animals studies.
  - Liver tiles (299 × 299 pixels)
    - ballooning degeneration (0–2), lobular inflammation (0–3), steatosis (0–3) and fibrosis (0–4) [Kleiner score system]
- **Liver-NAS**
  - private dataset of liver biopsy images collecting from 9 patients.
  - whole slide images (~ 106259 × 306939 pixels)

  → Image tiles 224 × 224 pixels (to sufficiently identify the relevant histological features within the tile)

  ◆ Steatosis脂肪变性 (N = 3838),

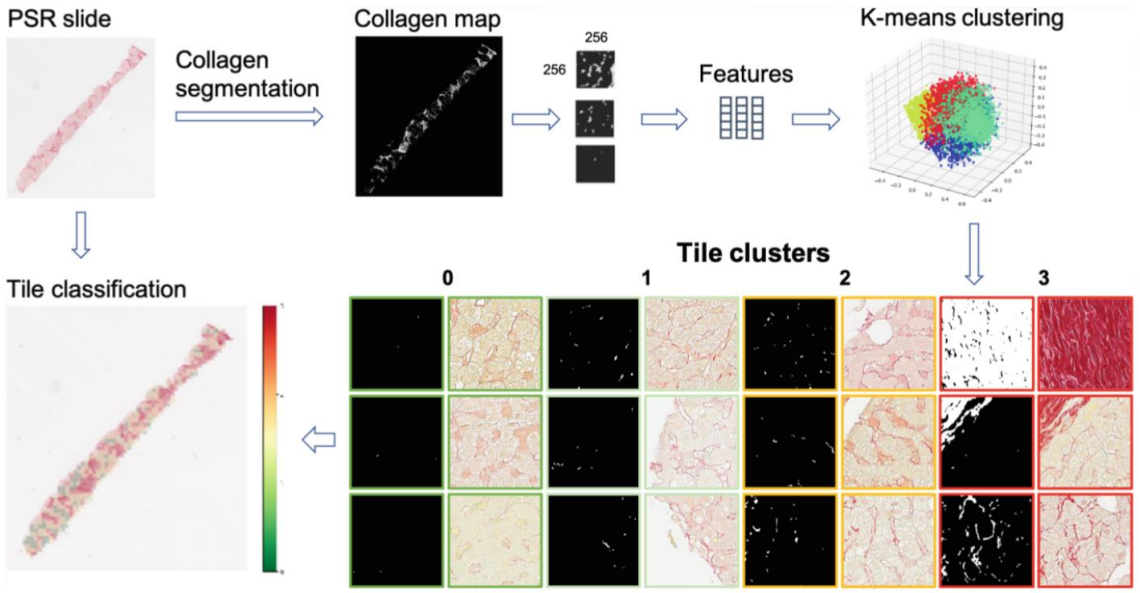
  ◆ ballooning degeneration气球样变性(N = 298),

  ◆ lobular inflammation (N = 69)小叶炎症

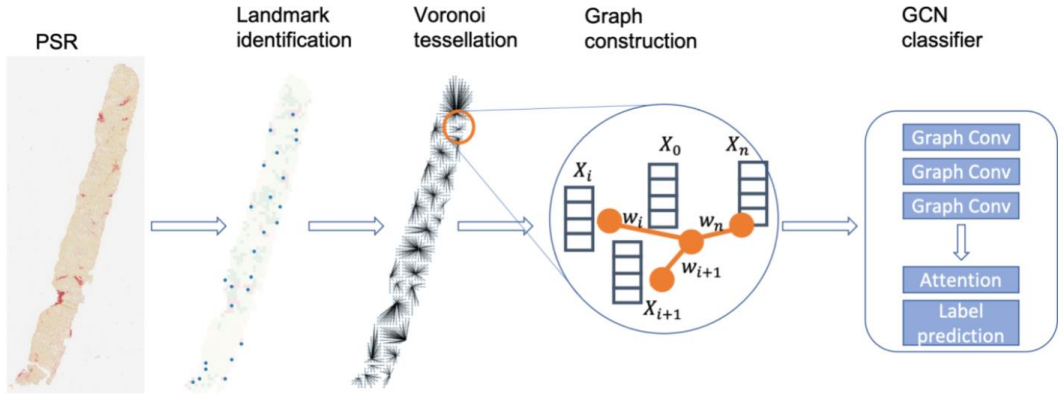
  ◆ others (N = 1659)
- Randomly split the data and report the results using 5-fold cross-validation
  - training set (70%), a validation set (10%) and test set (20%).

Methods

Early Detection of Liver Fibrosis Using Graph Convolutional Networks



Tile subtyping pipeline



The slide classification pipeline

Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification.

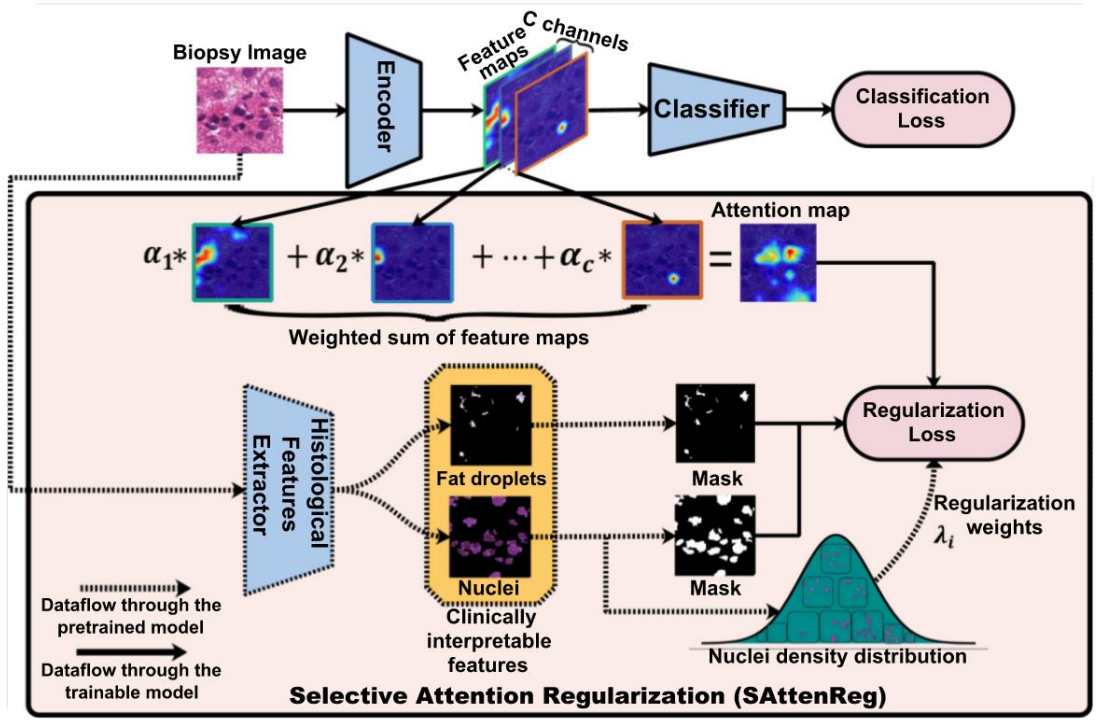


Fig. 1. Illustration of our proposed method.

临床可解释特征
- 细胞核 Nuclei
- 脂肪滴 Fat droplets

↓

guide and regularize attention maps → large overlap with the masks of clinically interpretable features.

Methods

Early Detection of Liver Fibrosis Using Graph Convolutional Networks

- Tile subtyping
  - CNN: identify regions that contain collagen.
    - Features from small tiles (256 × 256 px) by ImageNet pretrained ResNet18 model
    - resulting feature vectors$\in R^{512}$ are clustered using a k-means clustering algorithm. (k=4)
    - to learn the fibre representation.
  - Tiles from cluster 3.
- Graph Construction
  - regions of high collagen content are identified
  - centroids of each dense collagen region are used as centre points for Voronoi tessellation.
  - Tissue graph:
    - feature from each tile as node features X.
    - Euclidean distance D from each tile to its corresponding Voronoi centre is encoded as edge weight w.
- Graph Convolutional Layers
  - Input: Associated feature $X \in \mathbb{R}^{N,512}$; Adjacency matrix $A \in \mathbb{R}^{N,N}$.
  - 3-layer GNN:
    $H' = ReLu\big(G^*(X,A)\big), H'' = ReLu\big(G^*(H',A)\big), H = \tanh\big(G^*(H'',A)\big)$.
- Attention Layer
  - Aggregate the node vectors into a set of 4 vectors, each vector being a representation of the input slide with respect to one of the output classes.
  - attention for a class c:

$$v_c = \sum_i \alpha_{ci}\mathbf{H}_i \quad \text{with} \quad \alpha_c = Softmax(\mathbf{H}u_c^\top).$$

$$\hat{y} = Softmax(\hat{y}_0, ..., \hat{y}_C) \quad \text{with} \quad \hat{y}_c = \mathbf{W}_c^\top v_c + b_c.$$

Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification.

- Image Classification
  - Classificatio loss: $L_{cls} = \frac{1}{N}\sum_{i=1}^{N} y_i \log \hat{y}_i$
  - Predicated label: $\hat{y}_i = g \circ f(x_i)$
- Selective Attention Regularization
  - weighted sum of feature maps reflects the class-specific image regions used by the model for prediction

$$M_{atten}^i = \sum_{c=1}^{C} \alpha_c f_c(x_i)$$

  - For each input x_i, histological features extractor $\{E_{nuclei}, E_{fat}\}$ to get the nuclei mask M_{nuclei}^i and fat droplets mask M_{fat}^i:
    - $M_{nuclei}^i = E_{nuclei}(x_i);\ M_{fat}^i = E_{fat}(x_i)$
  - Regularization loss

    $L_{reg} = \frac{1}{N}\sum_{i=1}^{N} \lambda_i\, Dice\big(M_{nuclei}^i, M_{atten}^i\big) + (1-\lambda_i) Dice\big(M_{fat}^i, M_{atten}^i\big)$
- Make nuclei density conform to Gaussian: $\left|\left|M_{nuclei}^i\right|\right|_1 \sim N(\mu, \sigma^2)$
- The attention regularization weights λ_i is drawn from the Gaussian distribution

$$\lambda_i = \begin{cases} 1 - \frac{1}{\sigma\sqrt{2\pi}} exp^{-\frac{1}{2}(\frac{||M_{nuclei}^i||_1 - \mu}{\sigma})^2} & ||M_{nuclei}^i||_1 \geq \mu \\ \frac{1}{\sigma\sqrt{2\pi}} exp^{-\frac{1}{2}(\frac{||M_{nuclei}^i||_1 - \mu}{\sigma})^2} & ||M_{nuclei}^i||_1 < \mu \end{cases}$$

- $L = L_{cls} + L_{reg}$

Experiments

Early Detection of Liver Fibrosis Using Graph Convolutional Networks

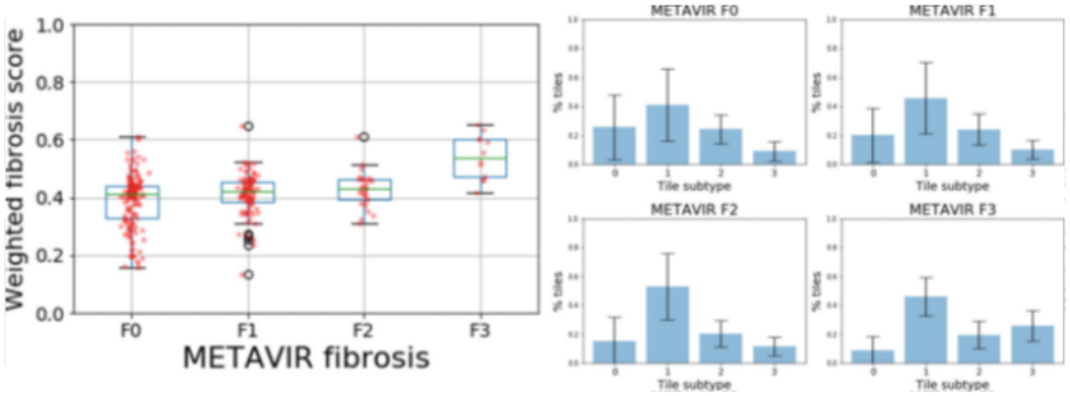


Fig. 4. The weighted fibrosis score across fibrosis stages and the distribution of tile subtypes across each METAVIR stage. The weighted fibrosis score, which is not informed by tissue topology, does not separate between classes F0 and F1 (Wilcoxon signed-rank test coefficients: W = −1.798, p = 0.072) and classes F1 and F2 (W = −0.775, p = 0.438). However, class F3 can be differentiated from all the other classes: F0 (W = −4.436, $p < 0.001$), F1 (W = −4.181, $p < 0.001$), F2 (W = −3.559, $p < 0.001$), and class F2 can be differentiated from F0 (W = −2.062, p = 0.039).

$$WFS = \frac{1}{C-1}\sum_{c=0}^{C-1} cp_c,$$

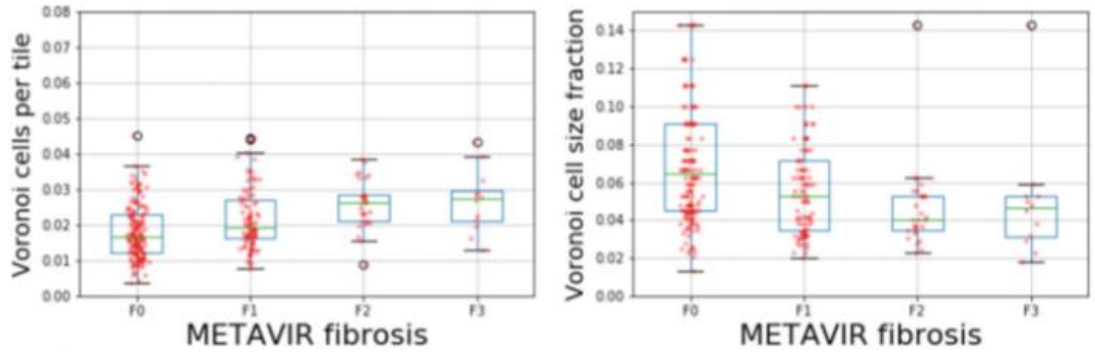
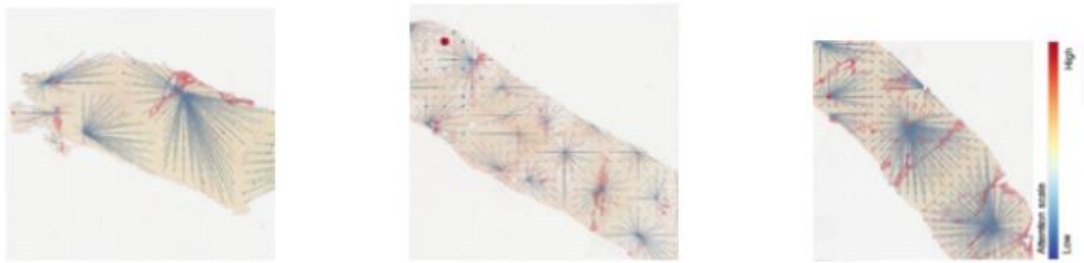


Fig. 5. The number and size of the Voronoi cells across the fibrosis stages. It can be seen that the size of a Voronoi cell relative to the area of the biopsy can change by an order of magnitude, therefore allowing to capture biological objects of varying sizes.



(a) (b) (c)

Fig. 6. Examples of tissue graphs superimposed on the original images. Node size and colour reflect attention activation for the predicted slide class. Notice that the edge length in the individual sub-graphs corresponds to the underlying pattern of fibrosis and that the sub-graphs capture biologically meaningful regions, e.g. (6a) Stage F0 with no signs of fibrosis. (6b) Stage F2 with early fibrotic clusters in the parenchyma. (6c) Stage F3 with prominent bridging.

Experiments

Early Detection of Liver Fibrosis Using Graph Convolutional Networks

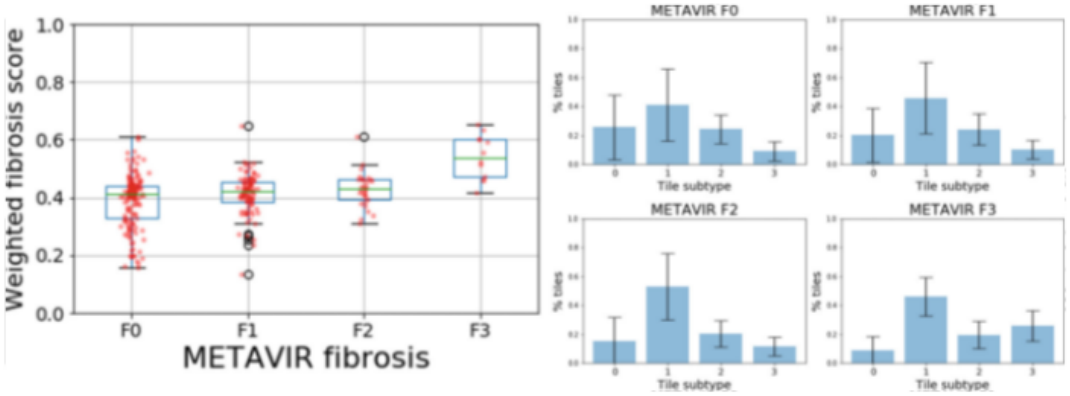


Fig. 4. The weighted fibrosis score across fibrosis stages and the distribution of tile subtypes across each METAVIR stage. The weighted fibrosis score, which is not informed by tissue topology, does not separate between classes F0 and F1 (Wilcoxon signed-rank test coefficients: W = −1.798, p = 0.072) and classes F1 and F2 (W = −0.775, p = 0.438). However, class F3 can be differentiated from all the other classes: F0 (W = −4.436, $p < 0.001$), F1 (W = −4.181, $p < 0.001$), F2 (W = −3.559, $p < 0.001$), and class F2 can be differentiated from F0 (W = −2.062, p = 0.039).

$$WFS = \frac{1}{C-1}\sum_{c=0}^{C-1} cp_c,$$

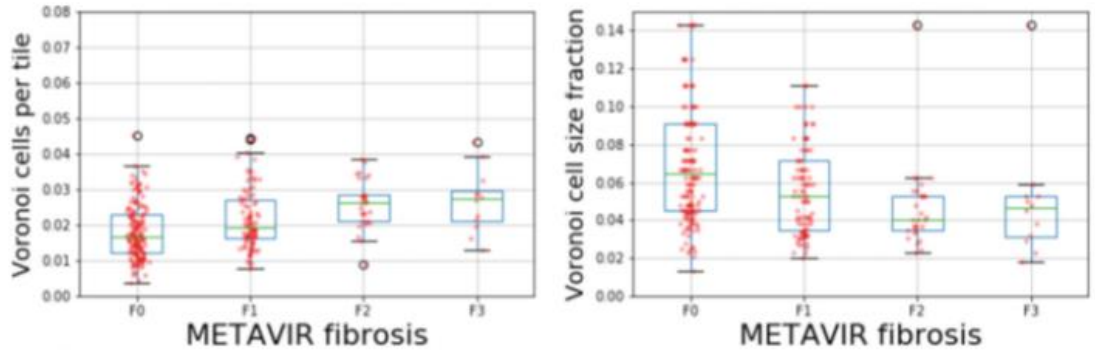
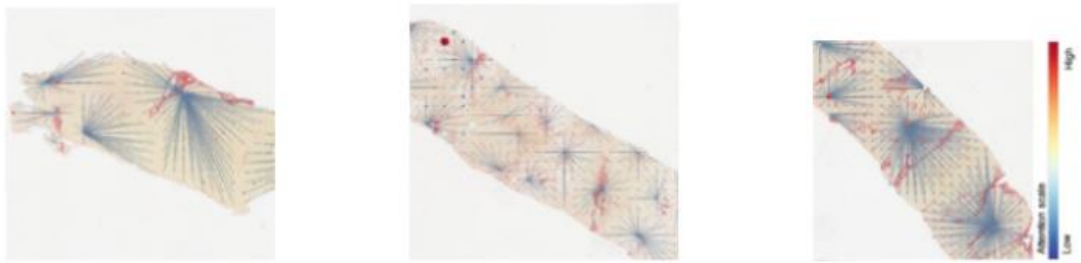


Fig. 5. The number and size of the Voronoi cells across the fibrosis stages. It can be seen that the size of a Voronoi cell relative to the area of the biopsy can change by an order of magnitude, therefore allowing to capture biological objects of varying sizes.



(a) (b) (c)

Fig. 6. Examples of tissue graphs superimposed on the original images. Node size and colour reflect attention activation for the predicted slide class. Notice that the edge length in the individual sub-graphs corresponds to the underlying pattern of fibrosis and that the sub-graphs capture biologically meaningful regions, e.g. (6a) Stage F0 with no signs of fibrosis. (6b) Stage F2 with early fibrotic clusters in the parenchyma. (6c) Stage F3 with prominent bridging.

Experiments

Early Detection of Liver Fibrosis Using Graph Convolutional Networks

Table 1. F1 scores across fibrosis classes. Results expressed in %.

Model	F0	F1	F2	F3
ResNet18	53.81 ± 6.36	21.40 ± 12.73	16.93 ± 15.00	24.44 ± 21.43
GCN	65.09 ± 5.50	19.75 ± 11.63	**28.15** ± 24.48	70.71 ± 12.02
ATN-GCN	51.63 ± 2.83	27.27 ± 15.62	21.99 ± 6.72	**75.56** ± 7.70
GAT	**71.32** ± 8.56	**34.56** ± 21.11	0.00 ± 0.00	44.44 ± 38.49
ATN-GAT	59.15 ± 9.09	33.23 ± 28.91	22.22 ± 38.49	19.05 ± 16.49
GIN	68.31 ± 0.23	13.33 ± 23.09	0.00 ± 0.00	13.33 ± 23.09
ATN-GIN	68.18 ± 0.00	0.00 ± 0.00	0.00 ± 0.00	0.003 ± 0.00
	F0–F1	–	F2	F3
ResNet18	67.98 ± 2.26		26.07 ± 7.74	0.00 ± 0.00
GCN	**93.01** ± 0.53		18.79 ± 18.21	0.80 ± 0.00
ATN-GCN	89.26 ± 2.58		**28.28** ± 11.00	**78.57** ± 6.18
GAT	92.07 ± 1.40		27.78 ± 25.46	61.11 ± 9.62
ATN-GAT	86.22 ± 6.98		0.00 ± 0.00	0.00 ± 0.00
GIN	90.02 ± 0.00		0.00 ± 0.00	0.00 ± 0.00
ATN-GIN	90.02 ± 0.00		0.00 ± 0.00	0.00 ± 0.00
	F0–F1–F2	–	–	F3
ResNet18	74.19 ± 1.78			6.36 ± 5.53
GCN	99.03 ± 0.00			85.71 ± 0.00
ATN-GCN	**99.36** ± 0.55			**88.57** ± 10.30
GAT	96.85 ± 1.08			38.89 ± 34.69
ATN-GAT	96.90 ± 1.04			0.00 ± 0.00
GIN	96.30 ± 0.00			0.00 ± 0.00
ATN-GIN	96.30 ± 0.00			0.00 ± 0.00

Experiments

Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification.

Table 1. Quantitative comparisons on the private Liver-NAS dataset for NAS-related components classification. ($\lambda_i = \mathcal{N}$ refers to Gaussian distribution-induced weights)

Methods		Others	LI	BD	Steatosis
Fabian et al. [7]	Sensitivity	91.14 ± 3.01	**43.82** ± 21.67	78.16 ± 15.97	97.06 ± 3.50
	Specificity	95.12 ± 2.52	99.38 ± 0.65	99.26 ± 7.81	95.86 ± 1.78
	F1	89.66 ± 4.09	45.46 ± 23.49	80.62 ± 6.61	97.38 ± 1.42
SCNet [13]	Sensitivity	93.34 ± 3.41	28.50 ± 19.07	75.90 ± 21.58	97.28 ± 2.18
	Specificity	95.10 ± 1.71	99.72 ± 0.43	99.08 ± 0.96	**96.74** ± 2.51
	F1	90.74 ± 3.59	37.60 ± 24.44	76.86 ± 12.52	97.72 ± 0.84
Ours ($\lambda_i = \frac{1}{2}$)	Sensitivity	93.54 ± 2.75	35.44 ± 15.87	76.68 ± 17.72	97.78 ± 1.82
	Specificity	95.80 ± 1.39	99.66 ± 0.24	99.22 ± 0.81	96.56 ± 1.99
	F1	91.64 ± 2.52	43.38 ± 17.97	79.12 ± 12.81	97.96 ± 0.58
Ours ($\lambda_i = \mathcal{N}$)	Sensitivity	**93.92** ± 1.17	43.48 ± 20.38	**78.58** ± 14.59	**97.94** ± 1.43
	Specificity	**95.96** ± 1.19	**99.74** ± 0.21	**99.42** ± 0.64	96.60 ± 1.53
	F1	**92.02** ± 1.89	**51.60** ± 21.93	**82.02** ± 6.35	**98.06** ± 0.46

Table 2. Quantitative comparisons for weakly supervised localization ability with or without SAttenReg on private Liver-NAS dataset.

Method	Dice		
	LI	BD	Steatosis
W/o SAttenReg	0.4299 ± 0.1065	0.3713 ± 0.1030	0.2262 ± 0.1128
With SAttenReg	**0.4699** ± 0.1034	**0.5072** ± 0.0640	**0.6144** ± 0.0375

Table 3. Quantitative comparisons on quantifying the NAS-related components on the Biopsy4Grading [7] dataset.

Methods		LI	BD	Steatosis
Fabian et al. [7]	Sensitivity	79.36 ± 2.27	85.83 ± 1.35	94.20 ± 1.50
	Specificity	94.60 ± 0.60	93.94 ± 0.78	98.73 ± 0.38
	F1	80.16 ± 1.63	87.61 ± 0.90	93.78 ± 1.16
SCNet [13]	Sensitivity	78.52 ± 2.68	85.63 ± 2.44	94.13 ± 1.38
	Specificity	94.31 ± 0.81	93.55 ± 1.04	98.68 ± 0.28
	F1	78.94 ± 1.83	87.11 ± 1.38	93.85 ± 0.98
Ours ($\lambda_i = \frac{1}{2}$)	Sensitivity	78.53 ± 2.72	86.30 ± 1.08	94.17 ± 1.08
	Specificity	94.39 ± 2.72	93.91 ± 0.57	98.70 ± 0.30
	F1	79.24 ± 2.49	87.90 ± 0.98	93.90 ± 0.79
Ours ($\lambda_i = \mathcal{N}$)	Sensitivity	**80.42** ± 2.17	**86.46** ± 2.93	**94.68** ± 1.12
	Specificity	**94.92** ± 0.99	**94.13** ± 1.15	**98.78** ± 0.29
	F1	**81.03** ± 0.99	**88.01** ± 1.84	**94.42** ± 0.81

Experiments

Focusing on Clinically Interpretable Features: Selective Attention Regularization for Liver Biopsy Image Classification.

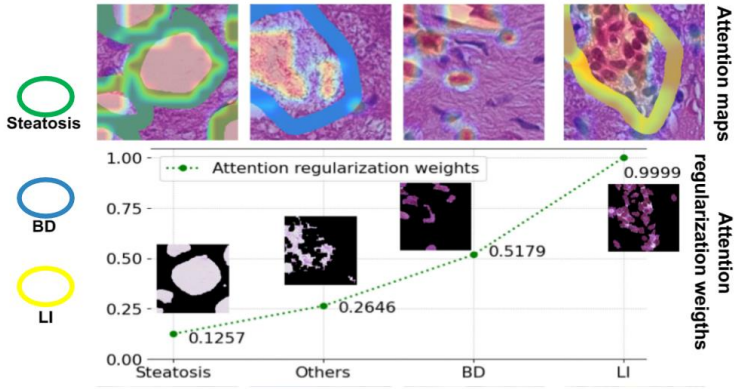


Fig. 2. Visual attention maps (upper row) and corresponding attention regularization weights (lower row) based on the nuclei density. The circle denotes the ground truth region (NAS-related components) annotated by pathologists (best viewed in color).

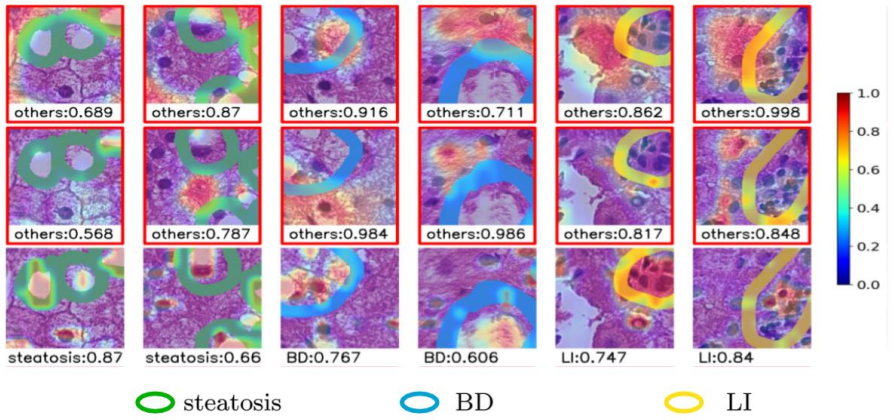


Fig. 3. Visual comparisons of attention maps generated by Grad-CAM [15] between Fabian et al. [7] (1st row), SCNet [13] (2nd row) and ours (3rd row). The circle denotes the ground truth region (NAS-related components) annotated by pathologists. Red bounding box indicates the false predication (best viewed in color). The predicted class and probability are shown at the bottom of the picture.