

Fig 1. Sample distribution of different cancer pathology images and schematic diagram of laryngeal cancer pathology images. Inter-cellular bridge (IB), atypical cells (AC), more chaotic arrangement (MCA), and darker staining (DS). Existing methods based on the transfer of natural images to pathological images result in poor transfer effects because there are differences in the task requirements and data samples. However, transfer screening work is more arduous with the method of migration from similar cell images due to the lack of prior experience and analysis of sample distribution.

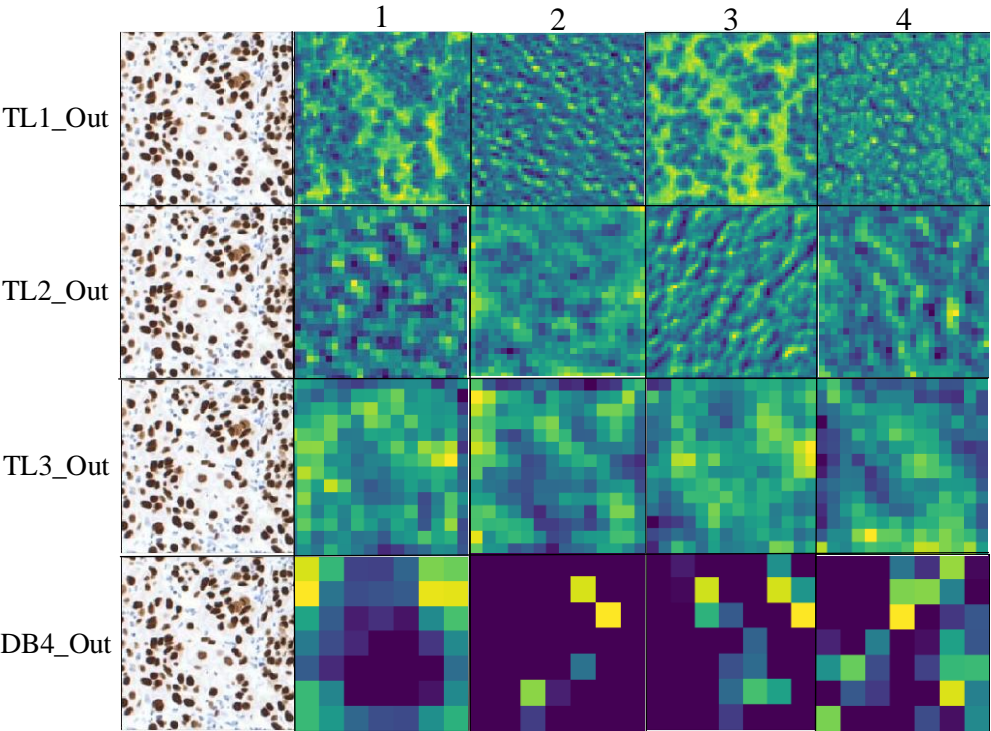


Fig 2. Visualization of feature channels with different insertion positions.

TABLE 1. Experimental results of model transfer learning based on a priori clinical experience and sample analysis.

Note: Lymph→Larynx indicates that the weights trained from the Lymph dataset are the initial weights of Larynx training.

Model Method	Resnet50[36]	DenseNet121[26]	Inception[37]	Xception[38]	VGG19[39]
Larynx[1]	42.35%	51.76%	44.11%	42.94%	39.41%
Lymph[30]	89.55%	92.22%	86.00%	90.22%	61.11%
Cervix[33-34]	77.94%	82.97%	81.25%	82.28%	67.19%
Brain[31]	96.67%	98.33%	97.08%	96.25%	74.84%
Breast[35]	74.46%	75.74%	54.46%	77.87%	36.17%
Lymph→Larynx	74.70% ↑32%	77.64% ↑25%	74.70% ↑33%	74.70% ↑32%	55.88% ↑16%
Cervix→Larynx	58.23% ↑16%	72.94% ↑21%	65.88% ↑22%	71.17% ↑28%	38.82% ↓0.6%
Brain→Larynx	67.64% ↑25%	72.35% ↑20%	70.58% ↑26%	73.52% ↑31%	48.82% ↑9.0%
Breast→Larynx	72.94% ↑30%	69.99% ↑18%	60.00% ↑16%	66.47% ↑24%	48.23% ↑9.0%
ImageNet→Larynx	68.82% ↑26%	67.05% ↑16%	72.94% ↑28%	72.94% ↑30%	41.76% ↑2.3%

TABLE 2. Experimental results of FAB module. Note: FAB 2-, FAB 3-, and FAB 4- mean to insert the FAB module after TL2, TL3, and DB4, respectively, of DenseNet121; please refer to Section III.

Fusion Method Insertion method	AFABNet	CFABNet	SFABNet	SENet[27]
FAB 1	73.52% ↓4%	73.52% ↓4%	68.82% ↓9%	74.47% ↓3%
FAB 2	85.29% ↑7%	85.29% ↑7%	68.23% ↓9%	72.35% ↓5%
FAB 3	77.64% ↑0%	83.52% ↑6%	65.29% ↓12%	73.52% ↓4%
FAB 4	77.64% ↑0%	75.88% ↓2%	80.00% ↑2%	72.94% ↓5%
FAB 1-FAB 2	77.05% ↓1%	85.29% ↑8%	71.76% ↓6%	75.29% ↓2%
FAB 1- FAB 3	77.05% ↓1%	79.41% ↑2%	68.82% ↓9%	75.88% ↓1%
FAB 1- FAB 4	80.58% ↑3%	78.82% ↑1%	71.17% ↓6%	72.35% ↓5%
FAB 2- FAB 3	86.47% ↑9%	77.64% ↑0%	65.29% ↓12%	74.11% ↓3%
FAB 2- FAB 4	87.65% ↑10%	81.76% ↑4%	75.29% ↓2%	69.99% ↓7%
FAB 3- FAB 4	79.41% ↑2%	82.35% ↑5%	74.11% ↓3%	75.88% ↓1%
FAB 1- FAB 2- FAB 3	78.23% ↑1%	82.35% ↑5%	65.88% ↓12%	70.58% ↓7%
FAB 1- FAB 2- FAB 4	72.35% ↓5%	69.41% ↓8%	70.58% ↓7%	75.29% ↓2%
FAB 1- FAB 3- FAB 4	85.29% ↑8%	70.58% ↓7%	77.11% ↑0%	71.64% ↓6%
FAB 2- FAB 3- FAB 4	85.88% ↑8%	82.94% ↑5%	65.88% ↓12%	73.52% ↓4%
FAB 1-FAB 2-FAB 3-FAB 4	85.88% ↑8%	84.70% ↑7%	57.64% ↓20%	68.23% ↓9%

$$p_{j|i} = \frac{\exp(-\|x_i - x_j\|^2 / 2\sigma^2)}{\sum_{k \neq i} \exp(-\|x_i - x_k\|^2 / 2\sigma^2)} \quad (1)$$

In formula (1), $p_{j|i}$ denotes the conditional probability of Euclidean distance between high-dimensional sample points, x_i denotes a high-dimensional sample point, x_j denotes a nearest neighbor of x_i in the high-dimensional data, and σ is the standard deviation of the Gaussian distribution.

$$q_{j|i} = \frac{\exp(-\|y_i - y_j\|^2)}{\sum_{k \neq i} \exp(-\|y_i - y_k\|^2)} \quad (2)$$

In formula (2), $q_{j|i}$ denotes the conditional probability of the Euclidean distance between the low-dimensional sample points, y_i denotes the low-dimensional sample points, and y_j is denoted as the nearest neighbor of x_i in the low-dimensional data.

$$C = \sum_i KL(P_i || Q_i) = \sum_i \sum_j p_{j|i} \log \frac{p_{j|i}}{q_{j|i}} \quad (3)$$

In formula (3), C minimizes the KL scatter between two distributions (Kullback-Leibler divergence).