1. Kim, E., Kim, S., Seo, M., & Yoon, S. (2021). **XProtoNet: Diagnosis in Chest Radiography with Global and Local Explanations.** In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 15719-15728).
2. Zhang, H., Zhang, J., Wang, R., Zhang, Q., Gauthier, S. A., Spincemaille, P., ... & Wang, Y. (2021, April). **Geometric Loss for Deep Multiple Sclerosis lesion Segmentation.** In 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) (pp. 24-28). IEEE. (Best Paper Finalist

Gao, Lufei

2021.7.7

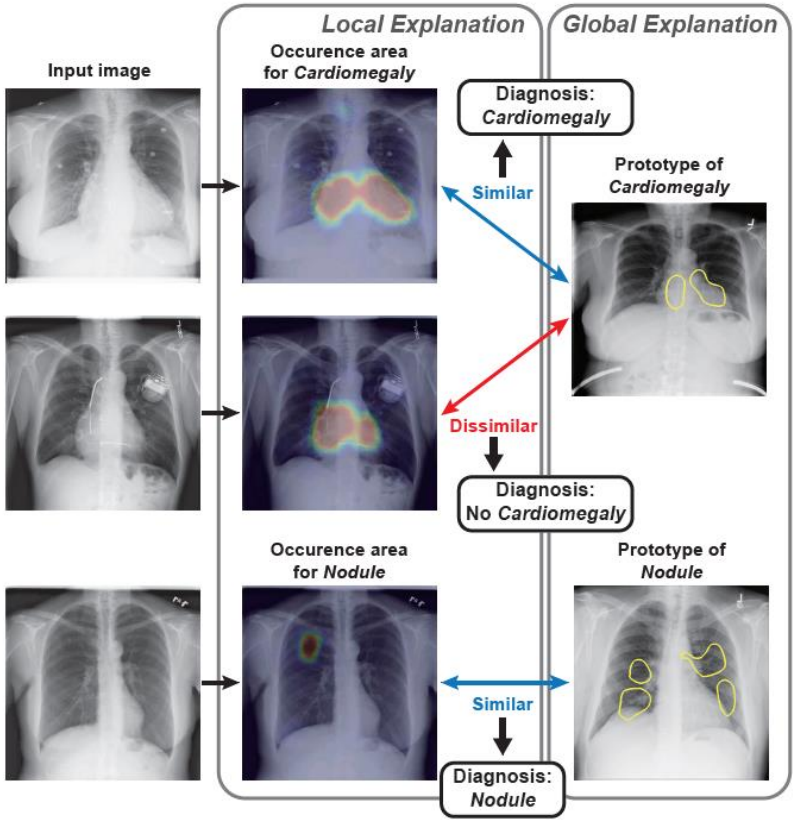
XProtoNet: Diagnosis in Chest Radiography with Global and Local Explanations.

- 医疗诊断应用不仅需要高性能，而且需要有强有力的判断依据。
- 许多自动诊断方法都将定位作为预测的解释，但这只提供了给定图像中网络聚焦的区域，而不是网络做出决策的方式。
- Interpretable models （可解释模型）
  - Case-based models适合于构建可解释的自动诊断系统。
    - 学习每一类的discriminative features – prototypes
    - 将输入图像的特征与prototypes进行比较进行分类。
    - 两种解释:全局解释global explanation和局部解释local explanation

Global	a class-representative feature that is shared by multiple data points belonging to the same class	定义每个类的共同特征	放射科医生的方式解释x射线图像的常见疾病的迹象
local	shows how the prediction of a single input image is made	找到模型将给定输入图像归入特定类的原因	通过检查给定的X 光图像中提供某种疾病信息的部分来诊断个别病例的方式

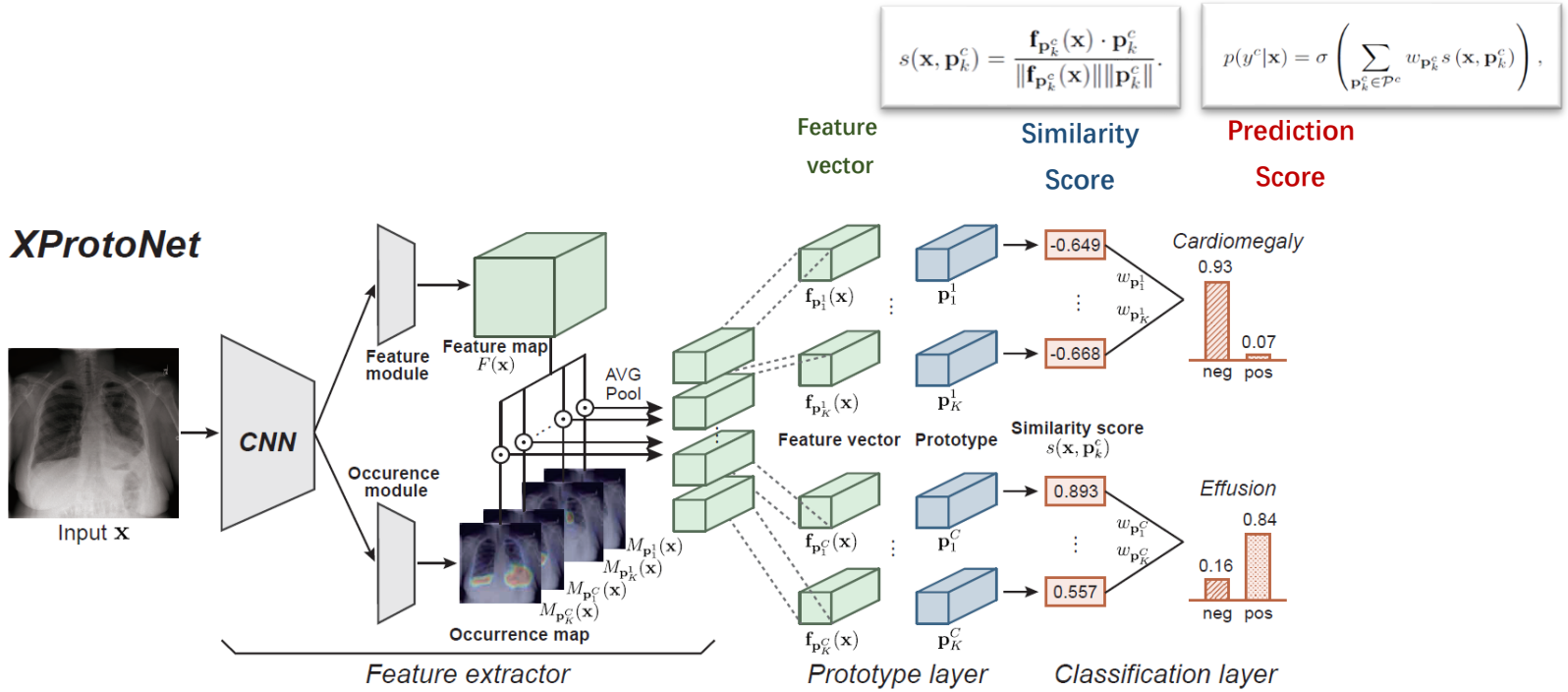
ProtoPNet → XProtoNet

- ProtoPNet
  - Prototype: a feature within a patch of a predefined size obtained from training images. 从训练图像中获得的预定义大小patch中的特征
  - compares a local area in a given input image with the prototypes for classification. 将给定输入图像中的局部区域与原型进行比较从而分类。
  - Constraint: with a patch of predefined size，prototypes
    - 很难反映动态区域中出现的特征 features that appear in a dynamic area
    - 无法充分呈现累代表特征 class-representative feature
    - 呈现与类无关的特征 a class-irrelevant feature。
- XProtoNet: present class-representative features within a dynamic area
  - 首个可解释的胸片诊断模型，可提供全局和局部解释。
  - 一种在动态区域内学习疾病代表特征的新方法，提高了可解释性和诊断性能。
  - 证明XProtoNet在公共NIH胸部x射线数据集上优于其他最先进的方法。



Overall architecture of XProtoNet

- A set of K learned prototypes $\mathcal{P}^c = \{p_k^c\}_{k=1}^K$ for each disease c
- p_k^c: a discriminative feature of disease c.
- s: 输入图像特征与每个prototype的相似程度
- $w_{p_k^c}$: 每个prototype对诊断的重要程度
- 训练结束后，用训练图像中最相似的特征向量$f_{p_k^c}$替换原型p_k^c。这使得prototype可以可视化为人类可解释的训练图像，



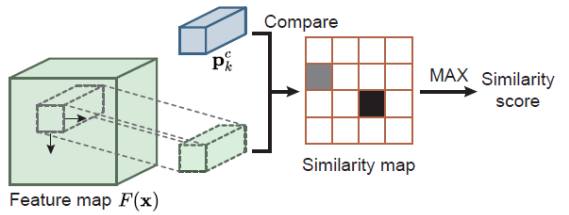
XProtoNet: Extraction of Feature with Occurrence Map

- 提取特征向量时，考虑输入图像的两个方面：
  - The patterns within the image
  - The area on which to focus to identify a certain disease.
- 特征提取对以上两个方面各包括：a feature module (特征模块) and an occurrence module (发生模块).
  - Feature module:
    - 提取Feature map $F(x) \in \mathbb{R}^{H \times W \times D}$: latent representation of input x
  - Occurrence module:
    - 对每个prototype p_k^c 预测Occurrence map $M_{p_k^c}(x) \in \mathbb{R}^{H \times W}$: 对应prototype可能出现的位置（焦点区域，focus area）
- 特征向量由两个模块得到（$u \in [0, H \times W]$ 表示空间位置）

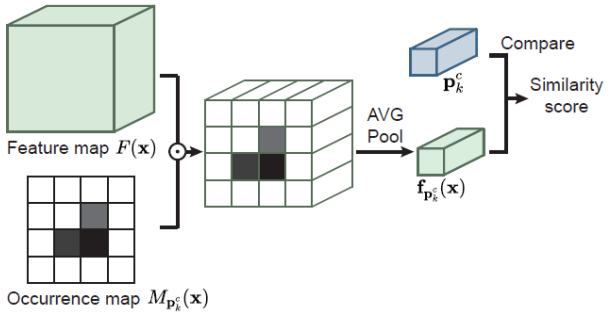
$$\mathbf{f}_{\mathbf{p}_k^c}(\mathbf{x}) = \sum_u M_{\mathbf{p}_k^c,u}(\mathbf{x}) F_u(\mathbf{x}),$$

- 池化feature map F(x)时，使用occurrence map的值作为权值，使得特征向量表示occurrence map高度活跃区域的一个特征。

ProtoPNet



XProtoNet



Training Scheme

- Classification loss

$$\mathcal{L}_{\text{cls}}^{c} = -\sum_{i} \frac{1}{|N_{\text{pos}}^{c}|}(1-p_i^c)^{\gamma} y_i^c log(p_i^c) - \sum_{i} \frac{1}{|N_{\text{neg}}^{c}|}(p_i^c)^{\gamma}(1-y_i^c) log(1-p_i^c),$$

- Regularization for Interpretability
  - 阳性样本x和prototype p_k^c之间的相似性应该大，阴性样本的相似性应该小
  - 聚类损失L_{lst}^c最大化阳性样本的相似度，分离损失L_{sep}^c最小化阴性样本的相似度

$$\mathcal{L}_{\text{clst}}^{c} = -y^c \max_{\mathbf{p}_k^c \in \mathcal{P}^c} s(\mathbf{x}, \mathbf{p}_k^c),$$

$$\mathcal{L}_{\text{sep}}^{c} = (1-y^c) \max_{\mathbf{p}_k^c \in \mathcal{P}^c} s(\mathbf{x}, \mathbf{p}_k^c).$$

- Regularization for Occurrence Map
  - Transformation loss（图像的仿射变换不会改变疾病迹象的相对位置）
  - Occurrence loss (L1 loss实现局部化，避免覆盖过多区域)

$$\mathcal{L}_{\text{trans}}^{c} = \sum_{\mathbf{p}_k^c \in \mathcal{P}^c} \|A(M_{\mathbf{p}_k^c}(\mathbf{x})) - M_{\mathbf{p}_k^c}(A(\mathbf{x}))\|_1,$$

$$\mathcal{L}_{\text{occur}}^{c} = \mathcal{L}_{\text{trans}}^{c} + \sum_{\mathbf{p}_k^c \in \mathcal{P}^c} \|M_{\mathbf{p}_k^c}(\mathbf{x})\|_1.$$

- Overall Cost Function

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{cls}} + \lambda_{\text{clst}}\mathcal{L}_{\text{clst}} + \lambda_{\text{sep}}\mathcal{L}_{\text{sep}} + \lambda_{\text{occur}}\mathcal{L}_{\text{occur}},$$

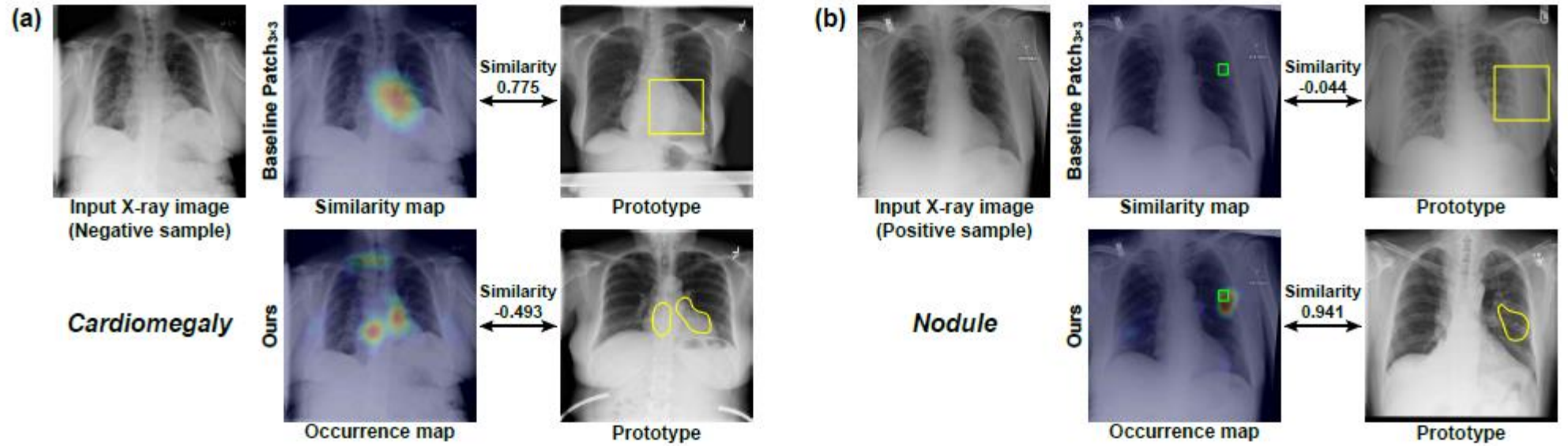
Experiments

- public NIH chest X-ray dataset
  - 112,120 frontal-view X-ray images with 14 disease labels from 30,805 unique patients.

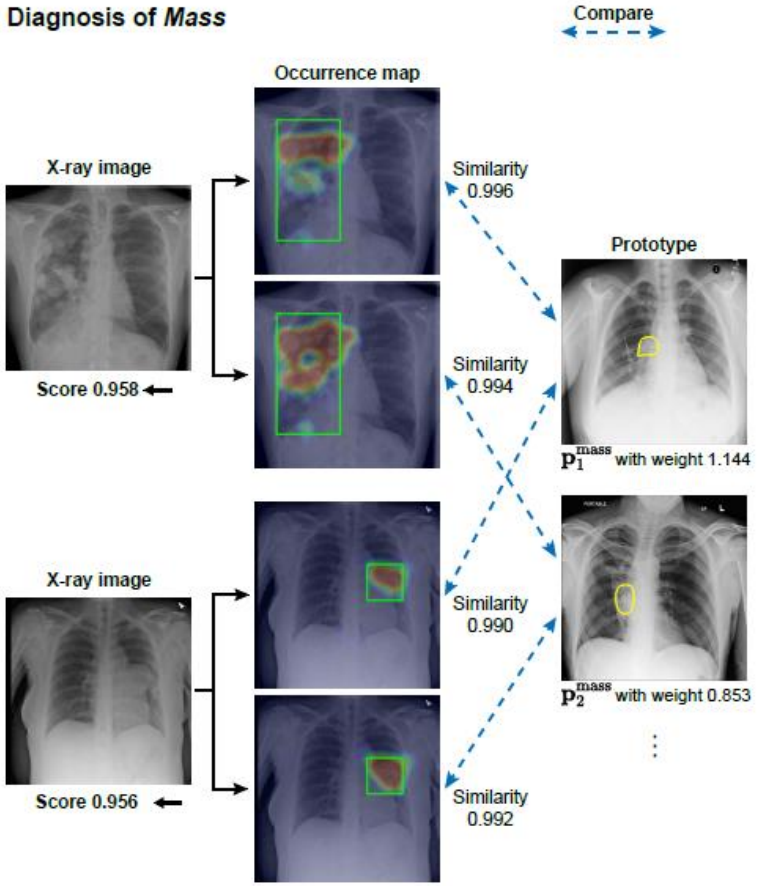
Table 1. AUC scores of XProtoNet and various baselines on chest X-ray dataset. The 14 diseases are Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Consolidation, Edema, Emphysema, Fibrosis, Pleural Thickening, and Hernia, respectively. The name of each disease is shortened to the first four characters (e.g. Atelectasis to Atel). Pne1, Pne2, and P.T. denote Pneumonia, Pneumothorax, and Pleural Thickening, respectively. The term "w/o $\mathcal{L}_{trans}$" denotes XProtoNet trained without $\mathcal{L}_{trans}$.

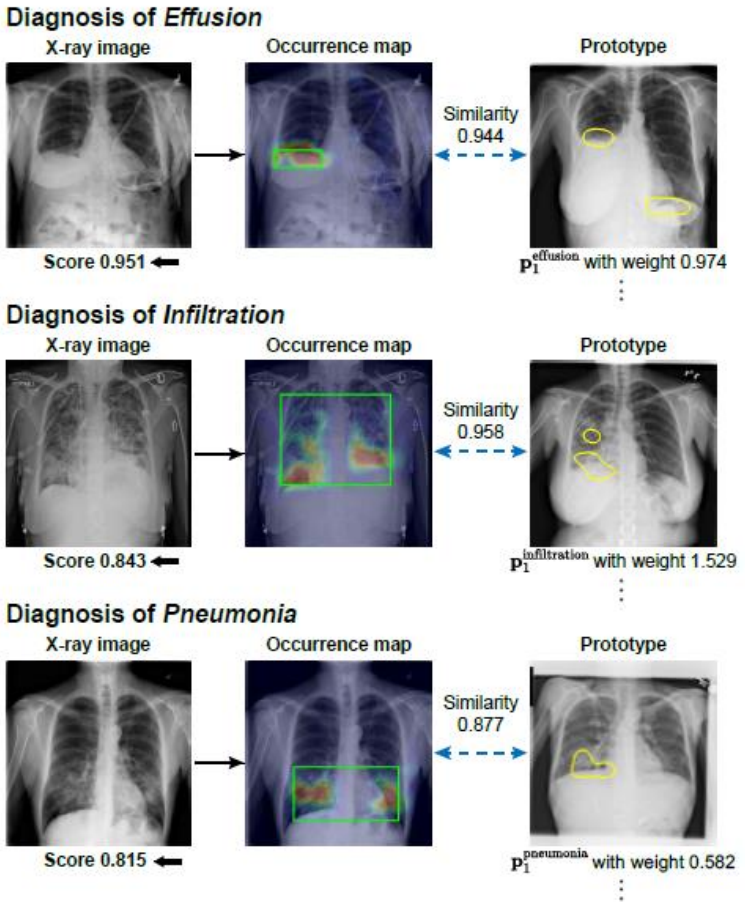
Methods	Atel	Card	Effu	Infi	Mass	Nodu	Pne1	Pne2	Cons	Edem	Emph	Fibr	P.T.	Hern	Mean
Baseline $Patch_{1\times1}$	0.766	0.857	0.823	0.705	0.813	0.779	0.706	0.851	0.738	0.825	0.925	0.779	0.771	0.663	0.786
Baseline $Patch_{3\times3}$	0.767	0.853	0.826	0.706	0.813	0.786	0.705	0.861	0.737	0.827	0.927	0.782	0.776	0.714	0.792
Baseline $Patch_{5\times5}$	0.752	0.863	0.822	0.695	0.814	0.751	0.702	0.834	0.734	0.827	0.906	0.793	0.772	0.543	0.772
Baseline GAP	0.764	0.847	0.815	0.703	0.817	0.782	0.719	0.856	0.723	0.823	0.928	0.782	0.776	0.704	0.789
XProtoNet (Ours)	**0.782**	**0.881**	**0.836**	**0.715**	**0.834**	**0.799**	**0.730**	**0.874**	**0.747**	**0.834**	**0.936**	**0.815**	**0.798**	**0.896**	**0.820**
w/o $\mathcal{L}_{trans}$	0.777	0.875	0.833	0.703	0.828	0.795	0.726	0.871	**0.747**	0.832	0.934	0.806	0.796	0.892	0.815

ProtoPNet中，性能与Patch的大小有关。
GAP（特征向量没有用Occurrence map池化）与Patch相似。
由于XProtoNet预测了要进行比较的自适应区域，所以它在所有类别中都比基线取得了更高的性能 。



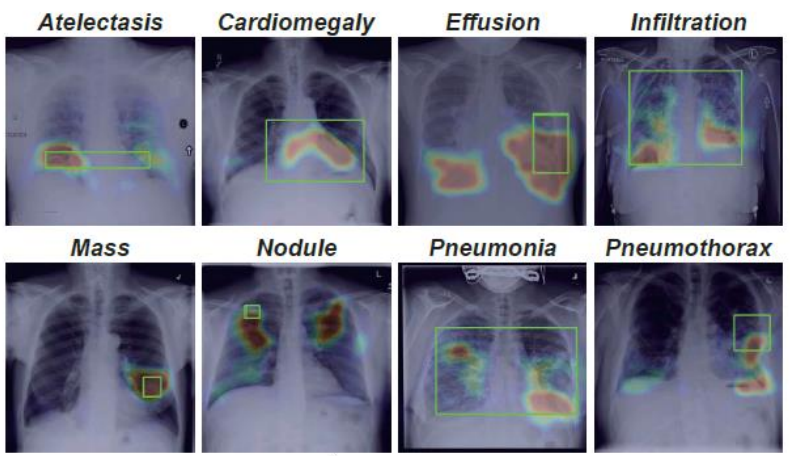
- Cardiomegaly（心脏肿大）
  - Baseline prototype只显示了心脏的一部分，与阴性样本相似性很高(0.775)。
  - XProtoNet的原型几乎可以显示整个心脏区域——更容易解释。两个发生区之间的相似度评分较低(-0.369)。
- Nodule （阳性结节）
  - XProtoNet成功检测到与原型相似度较高的小结节(0.936)
  - XProtoNet的结节原型对应的发生区域与ground-truth bounding box一致。
  - 显示了更多的可解释的可视化原型和更准确的预测比基线。





对输入x射线图像的解释显示1~2个具有最大权值的疾病原型。

- 黄色表示学习的原型;
- 绿色表示来自数据集的真实边界框。



图像中的occurrence area与每种疾病的实际症状(绿框)的轨迹一致。表明这些prototype经过了良好的训练，能够呈现适当的疾病代表性特征。

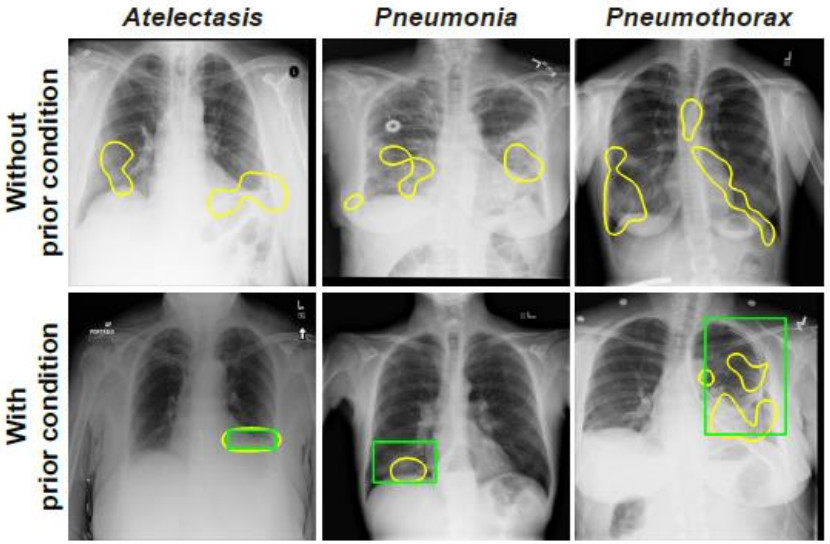
Table 2. AUC scores of XProtoNet and other methods on chest X-ray dataset. The * signifies that an additional conventional network is used as a backbone.

Methods	Atel	Card	Effu	Infi	Mass	Nodu	Pne1	Pne2	Cons	Edem	Emph	Fibr	P.T.	Hern	Mean
Backbone: ResNet-50															
Wang *et al.* [32]	0.700	0.810	0.759	0.661	0.693	0.669	0.658	0.799	0.703	0.805	0.833	0.786	0.684	0.872	0.745
Guan *et al.* [6]	0.779	0.879	0.824	0.694	0.831	0.766	0.726	0.858	**0.758**	**0.850**	0.909	**0.832**	0.778	**0.906**	0.814
XProtoNet (Ours)	**0.782**	**0.881**	**0.836**	**0.715**	**0.834**	**0.799**	**0.730**	**0.874**	0.747	0.834	**0.936**	0.815	**0.798**	0.896	**0.820**
Backbone: DenseNet-121 / DenseNet-121+α*															
Guan *et al.* [6]	**0.781**	0.883	0.831	0.697	0.830	0.764	0.725	0.866	**0.758**	**0.853**	0.911	0.826	0.780	0.918	0.816
Ma *et al.* [20]*	0.777	**0.894**	0.829	0.696	**0.838**	0.771	0.722	0.862	0.750	0.846	0.908	0.827	0.779	**0.934**	0.817
Hermoza *et al.* [10]*	0.775	0.881	0.831	0.695	0.826	0.789	**0.741**	**0.879**	0.747	0.846	0.936	**0.833**	0.793	0.917	0.821
XProtoNet (Ours)	0.780	0.887	**0.835**	**0.710**	0.831	**0.804**	0.734	0.871	0.747	0.840	**0.941**	0.815	**0.799**	0.909	**0.822**

Table 3. Comparison with methods that utilize additional bounding box annotations. AUC scores with a five-fold cross-validation performed on the chest X-ray dataset are reported. Following the previous works, the results are rounded to two decimal digits. The BBox column indicates whether bounding box annotation is used. Note that XProtoNet uses no additional supervision.

Methods	BBox	Atel	Card	Effu	Infi	Mass	Nodu	Pne1	Pne2	Cons	Edem	Emph	Fibr	P.T.	Hern	Mean
Li *et al.* [17]	✓	0.80	0.87	0.87	0.70	0.83	0.75	0.67	0.87	**0.80**	0.88	0.91	0.78	0.79	0.77	0.81
Liu *et al.* [19]	✓	0.79	0.87	0.88	0.69	0.81	0.73	0.75	0.89	0.79	**0.91**	0.93	0.80	0.80	**0.92**	0.83
XProtoNet (Ours)		**0.83**	**0.91**	**0.89**	**0.72**	**0.87**	**0.82**	**0.76**	**0.90**	**0.80**	0.90	**0.94**	**0.82**	**0.82**	**0.92**	**0.85**

- Analysis with the prior condition that the prototypes of XProtoNet should present the features within the bounding box annotations.
- To utilize the bounding box annotations during training:
  - extract the feature vectors from the feature maps within the bounding boxes as $f_{p_k^c}^{bbox}(x) = \sum_{u \in bbox} M_{p_k^c,u}(x) F_u(x)$,
    - $bbox$ denotes the spatial location inside the bounding box.
  - Change L1 loss on the occurrence map for the box-annotated data to $\sum_{p_k^c \in \mathcal{P}^c} \sum_{u \notin bbox} M_{p_k^c,u}(x)$ to suppress the area outside the bounding box from being activated in the occurrence map.
  - To enable the prototypes to present the features within the bounding boxes, the prototype vectors are replaced with their most similar feature vectors $f_{p_k^c}^{bbox}$ from the box-annotated data.



- 在诊断性能上没有显著差异
  - 在14种疾病的XProtoNet训练中，有和没有prior condition的平均AUC评分分别为0.850和0.849

Geometric Loss for Deep Multiple Sclerosis lesion Segmentation

- ISBI2021 Best Paper Finalist
- 多发性硬化症(Multiple Sclerosis, MS)病变只占脑容量的一小部分，并且在形状、大小和位置上存在异质性.
- Previous Methods:
  - Region-based loss functions: binary cross entropy (BCE) or Dice loss
  - Data imbalance problem(只有一小部分脑体素属于病变): weighted BCE and Tversky loss
  - Geometric information:改善医学图像分割
    - 附加的解码器体系结构来生成与损耗评估相关的几何特征图
    - based on shape- or boundary-aware loss function:在ground-truth或predicted probability map上执行几何变换。
      - The distance transformation mapping (DTM) is used in both boundary (BD) loss and Hausdorff distance (HD) loss
      - 其中变换映射中的每个体素表示它与ROI的最近边界之间的距离。
      - 边界感知的损失函数强制网络聚焦于病变ROI的表面，从而解决了病变和背景体素之间的巨大不平衡。
      - 可以很好地分割大的目标，在小MS病变上不够好。
- 提出了一个新的Geometric loss function,
  - 解决数据不平衡
  - 利用多发性病变的几何性质

Geometric Loss Formulation

- The proposed GEO loss combines volumetric and geometric correlations in a single module

$$\mathcal{L}_G = \frac{\sum_{v\in\Omega} \Theta(s_v, g_v)\Psi(s, g, v, \phi)}{\sum_{v\in\Omega} \Gamma(s_v, g_v)} \qquad (1)$$

  - Θ: measures the voxel-wise volumetric correlations between the ground-truth lesion mask and the output probability map 真实的病变掩码与输出概率图之间的体素级的体积相关性
  - Ψ: computes the voxel-wise geometric correlations between the ground-truth and the output map.计算ground-truth和输出概率图之间的体素级几何关联
  - $v = (v_x, v_y, v_z)$: the spatial position vector
  - s is the output probability map,
  - $s_v \in [0, 1]$ is the value of s at position v.
  - g is the ground-truth binary lesion mask,
  - $g_v \in \{0, 1\}$ is the value of g at position v.
  - ϕ定义在空间域Ω上的一个空间不变算子，捕获边缘和到边缘的距离等局部几何信息。
- GEO loss formula can be used in existing CNN models and allows flexible selection of loss functions for training.
- The widely used region-based BCE and Dice loss functions can be seen as special cases

Geometric Loss Instantiations

- 三维物体的边界面积比其体积小一个数量级，在边界空间进行损耗计算可以缓解病变与背景体素之间的较大不平衡，有利于小病变的分割。
  - BD[9]和HD损失[10] 重加权方案有利于对大的物体进行分割，但对小病灶分割有害
- computationally efficient **convolutional filters** such as **first- and second-order gradient operators** for edge enhancement in the loss functions.
- FOG loss（first-order gradient）

Let the output of ϕ be a three-element vector, where $\phi(s,v) = (\frac{\partial s}{\partial v_x}, \frac{\partial s}{\partial v_y}, \frac{\partial s}{\partial v_z})$, $\phi(g,v) = (\frac{\partial g}{\partial v_x}, \frac{\partial g}{\partial v_y}, \frac{\partial g}{\partial v_z})$; also, letting $\forall\ v \in \Omega$, $\Theta(s_v, g_v) = 1$, and $\Psi(s,g,v,\phi) = ||\phi(s,v) - \phi(g,v)||^2$, we can define the FOG loss as following:

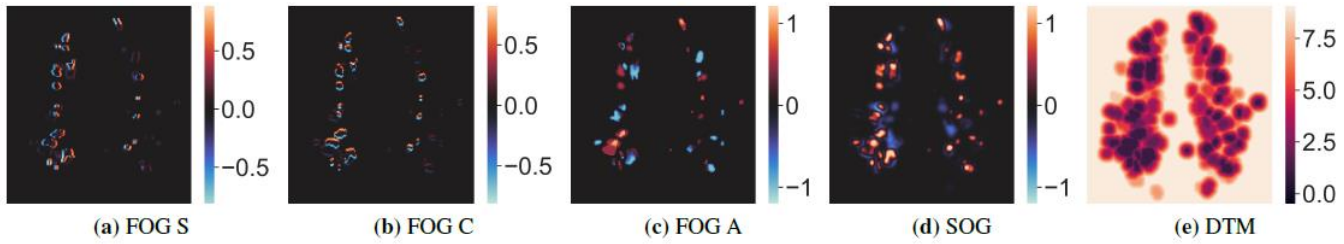
$$\mathcal{L}_F = \frac{1}{|\Omega|}\sum_{v\in\Omega} ||\phi(s,v) - \phi(g,v)||^2. \quad (2)$$

当追踪一个特定的病变时，神经放射学家通常检查所有轴向、矢状面和冠状面周围的切片。
进一步设计了三种FOG损耗，只计算其中一个正交平面上的梯度，其中$\phi(s,v)$=∂s/∂vi, $\phi(g,v)$ =∂g/∂vi, i表示{x,y,z}。

- SOG loss (second-order gradient)

The SOG loss is defined as the second-order differential operator which is the divergence of the gradient. Based on the boundary property of SOG, letting $\phi(s,v) = \frac{\partial^2 s}{\partial^2 v_x} + \frac{\partial^2 s}{\partial^2 v_y} + \frac{\partial^2 s}{\partial^2 v_z}$, $\phi(g,v) = \frac{\partial^2 g}{\partial^2 v_x} + \frac{\partial^2 g}{\partial^2 v_y} + \frac{\partial^2 g}{\partial^2 v_z}$, and $\Theta(s_v, g_v) = |s_v - g_v|$, SOG loss is derived as:

$$\mathcal{L}_S = \frac{1}{|\Omega|}\sum_{v\in\Omega} |s_v - g_v|\phi(g,v) \quad (3)$$



(a) FOG S (b) FOG C (c) FOG A (d) SOG (e) DTM

Experiments

- Two datasets with different scales were used for performance evaluation.
  - One small-scale dataset (GE-30) consisted of co-registered T1, T2, and T2-FLAIR images (0.7mm × 0.7mm × 3mm voxel size) acquired at a 3T GE scanner from 30 MS patients.
    - 方法对于训练样本有限的深度神经网络具有更强的鲁棒性。
  - Another large-scale dataset (SI-170) consisted of co-registered T1, T2, and T2-FLAIR images (1mm×1mm×1mm voxel size) acquired at a 3T SIEMENs scanner from 170 MS patients.
    - 方法在处理具有复杂背景细节的病灶分割方面也是有效的
  - The ground-truth masks of both datasets were traced by a neuroradiologist with 8 years of experience.

- Dice similarity coefficient (DSC), lesion-wise true positive rate (LTPR), lesion-wise positive predictive value (LPPV), and lesion-wise F1 score (L-F1) as evaluation metrics.

Table 1: Ablation study on the GE-30 dataset.

Methods	DSC	LPPV	LTPR	L-F1
Dice	0.705	0.580	0.834	0.684
Dice + FOG	0.712	**0.627**	0.863	**0.726**
Dice + FOG S	**0.715**	0.592	**0.898**	0.714
Dice + FOG C	0.708	0.602	0.889	0.718
Dice + FOG A	0.702	0.586	0.862	0.698
Dice + SOG One	0.708	0.606	0.857	0.710
Dice + SOG Two	0.706	0.600	0.855	0.705

Table 2: Performance comparison on the GE-30 dataset.

Methods	DSC	LPPV	LTPR	L-F1
LST [1]	0.594	0.418	0.708	0.526
Dice	0.705	0.580	0.834	0.684
Dice + HD [10]	0.502	0.393	0.608	0.477
Dice + BD [9]	0.704	0.594	**0.875**	0.708
Dice + FOG	0.712	**0.627**	0.863	**0.726**
Dice + SOG One	0.708	0.606	0.857	0.710

Table 3: Performance comparison on the SI-170 dataset.

Methods	DSC	LPPV	LTPR	L-F1
LST [1]	0.559	0.526	0.870	0.656
Dice	0.726	0.598	0.869	0.725
Dice + HD [10]	0.635	0.618	0.845	0.714
Dice + BD [9]	0.725	0.637	0.855	0.730
Dice + FOG	**0.737**	**0.666**	**0.873**	**0.756**
Dice + SOG One	0.735	0.664	0.860	0.749