DeGPR: Deep Guided Posterior Regularization For Multi-Class Cell Detection And Counting

Aayush Kumar Tyagi[1] Chiarg Mohapatra[1] Prasenjit Das[3] Govind Makharia[3], Lalita Mehra[3] Prathosh AP[2] Mausam[1]

[1]Indian Institute of Technology, Delhi, India

[2]Indian Institute of Science, Bangaluru, India

[3]All Indian Institute of Medical Science, Delhi, India

THU-PM-316 CVPR 2023

Outline

Motivation

- Detecting and counting cells in medical images : a crucial task
- Manual counting by pathologists:
  - Tiresome
  - Inter-observer inconsistencies
- Limited data, overlapping objects, multiple cell types, class imbalance and subtle variations in cell size and shape.
- How to guide a deep-learning model to look at the problem the way pathologists do?

Motivation

- DeGPR aims to bridge the gap between between deep learning object detectors and the pathologists that would use them by providing a system that can incorporate expert insights into object detectors.

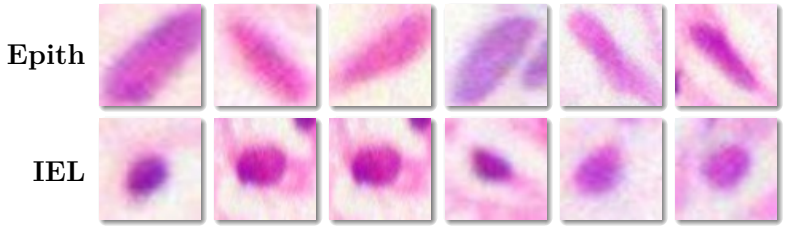


Figure: ENs are elongated, lightly stained while IELs are circular and dark stained.

Overview

- DeGPR enhances the object detector by explicitly considering the similarity between predicted and ground truth feature distributions. It incorporates two types of features:
  - **Explicit features**: Hand-crafted features, example: size and intensity
  - **Implicit features**:
    - Pathologits insights may not be complete.
    - Data driven and learnt directly from visual data.
    - Implicit features which are obtained from a pre-trained encoder.
- GMMs capture the feature distributions of ground truth and predicted bounding boxes. Then, it computes the KL divergence loss between the two GMMs.
- The DeGPR loss is added to the detection loss to jointly optimize object detection and feature distribution similarity.

DeGPR computation

Average Feature Value

$$\mathcal{A}_{f_j}(c) = \frac{1}{|B_c|} \sum_{b \in B_c} f_j(b) \tag{1}$$

Feature differences:

$$\mathcal{D}_{f_j}(c_i, c_k) = \mathcal{A}_{f_j}(c_i) - \mathcal{A}_{f_j}(c_k) \tag{2}$$

Feature Concatenation:

$$\mathcal{D}_F(c_i, c_k) = [\mathcal{D}_{f_1}(c_i, c_k);\ \mathcal{D}_{f_2}(c_i, c_k); \ldots\ ; \mathcal{D}_{f_m}(c_i, c_k)] \tag{3}$$

DeGPR computation

Fitting Gaussian Mixture model

- For each pair of classes:
  - Two separate GMMs, namely P_{gt} and Q_{pd} are trained.
- P_{gt} captures the statistical characteristics of feature differences in the ground truth bounding boxes
- Q_{pd} captures the statistical characteristics of feature differences in the predicted bounding boxes.
- The purpose of learning these GMMs is to align the feature vector distributions.

DeGPR computation

Compuatation of KL divergence loss:

$$D_{KL}(P_{gt}||Q_{pd}) = \int_{\mathcal{X}} P_{gt}(x) \ln \frac{P_{gt}(x)}{Q_{pd}(x)} dx \tag{4}$$

$$D_{MC} = \sum_{im} \log(P_{gt}(x_g)) - \log(Q_{pd}(x_p)) \tag{5}$$

Finally loss for all classes:

$$\mathcal{L}_F = \frac{1}{\binom{n}{2}} \sum_{i=1}^{n-1} \sum_{k=i+1}^{n} D_{KL}(P_{gt(i,k)}||Q_{pd(i,k)}) \tag{6}$$

DeGPR Features: Explicit Features

- Explicit features are based on pathologist insights
- Few of the features are:
  - Size of cell governed by size of bounding box.
  - Intensity of cell compute by average pixel intensity of bounding box.

DeGPR Features: Implicit features

- Implicit features are data driven and learned directly visual data using encoder.
- The encoder E_ϕ is pretrained using a supervised contrastive (SupCon) loss.
- The SupCon loss is hardness-aware, considering the difficulty of negative examples.
- The SupCon loss is computed as a sum over the set of all gold image patches.
- The loss aims to maximize similarity between positive pairs and minimize similarity between negative pairs.

DeGPR Features: Implict Features

Encoder Pre-trained:

$$L_{SupCon} = \sum_{v \in V} -\log \frac{1}{|P(v)|} \sum_{p \in P(v)} \frac{\exp(z_v z_p / \tau)}{\sum_a \exp(z_v z_a / \tau)} \tag{7}$$

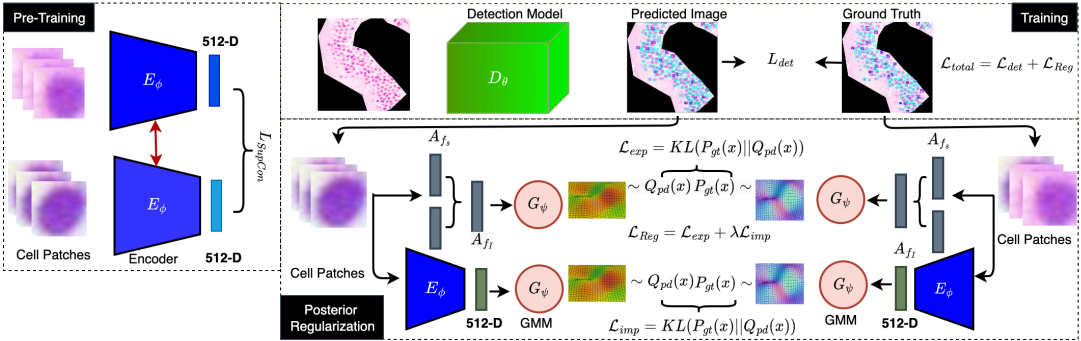
- Balanced sampling of patches per class is used during batch creation to address class imbalance.
- Gold patches are augmented by shifting and resizing bounding boxes to improve encoder robustness.
- Augmented bounding boxes are gradually introduced during training in an annealing manner.

Final loss

$$\mathcal{L}_{total} = \mathcal{L}_{det} + \lambda_{reg} \left(\mathcal{L}_{exp} + \mathcal{L}_{imp} \right) \tag{8}$$

- Combines the object detection loss with the regularization loss
- Optimizes both detection performance and feature alignment between ground truth and predicted bounding boxes.

Network Architecture



Results

Table: Detection and counting results for MuCeD

Model	Precision	Recall	mAP	MAE IEL	MRE IEL	MAE EN	MRE EN
Yolov5	0.711	0.723	0.751	8.97	42.62	14.61	13.43
Yolov5(DeGPR)	**0.744**	**0.735**	**0.787**	**5.83**	**24.19**	**13.15**	**12.46**
Faster-RCNN	0.592	0.436	0.496	11.85	50.05	27.50	24.93
Faster-RCNN (DeGPR)	**0.646**	**0.468**	**0.541**	**9.61**	**31.64**	**26.50**	**23.60**
EfficientDet	0.266	0.640	0.414	20.35	133.91	20.30	20.78
EfficientDet (DeGPR)	**0.274**	**0.641**	**0.425**	**17.32**	**90.04**	**18.51**	**18.12**

EN: Epithelial Nuclei

Results

Table: Detection and counting results for CoNSep

Model	Precision	Recall	mAP	MAE IF	MAE EN	MAE SP	MAE Avg
Yolov5	0.638	0.574	0.606	28.21	55.50	57.93	47.21
Yolov5 (DeGPR)	**0.667**	**0.584**	**0.625**	**26.35**	**55.00**	**53.85**	**45.07**
Faster-RCNN	0.490	0.208	0.342	64.71	227.93	198.29	163.64
Faster-RCNN (DeGPR)	**0.571**	**0.331**	**0.451**	**51.93**	**151.28**	**163.00**	**122.07**
EfficientDet	0.633	0.178	0.205	86.00	79.86	134.36	100.27
EfficientDet (DeGPR)	**0.672**	**0.194**	**0.229**	**79.64**	**77.78**	**125.85**	**94.42**

IF: Inflammatory cell, EN: Epithelial nuclei, SP: Spindle cell

Results

Table: Detection and counting results for MoNuSac

Model	Precision	Recall	mAP	MAE EN	MAE LY	MAE NP	MAE MP
Yolov5	0.611	**0.497**	0.481	25.15	14.12	1.96	3.95
Yolov5 (DeGPR)	**0.736**	0.474	**0.489**	**12.01**	**10.69**	**0.81**	**2.33**
Faster-RCNN	0.570	0.310	0.405	**19.52**	23.48	1.0	3.38
Faster-RCNN (DeGPR)	**0.643**	**0.370**	**0.473**	19.81	**22.44**	**0.82**	**3.02**
EfficientDet	0.256	**0.509**	0.402	17.67	17.25	1.24	6.51
EfficientDet (DeGPR)	**0.258**	0.499	**0.409**	**14.84**	**16.98**	**0.56**	**3.97**

EN:Epithelial nuclei, LY: Lymphocyte, NP: Neutrophil, MP: Macrophage

Results

Table: Counting vs Localization (MuCeD)

Model	MAE-IEL	MAE-Epith	MAE-Avg
UNet	11.72	26.85	19.29
FCRN-A	15.60	22.81	19.21
Countception	16.10	29.78	22.94
SAU-Net	11.56	28.07	19.82
Yolov5 (DeGPR)	**5.83**	**13.43**	**9.63**

Results

- Q-histology score is computed as the ratio of IELs per 100 ENs.
- If Q-histology score ≥ 25, then patient suffers from celiac disease.

Table: Classification Metrics based on Q-Ratio

Measure	Baseline	Yolo+DeGPR
Recall	0.774	**0.936**
Precision	0.774	**0.871**
F1-score	0.774	**0.902**
Accuracy	0.746	**0.877**

Visualization

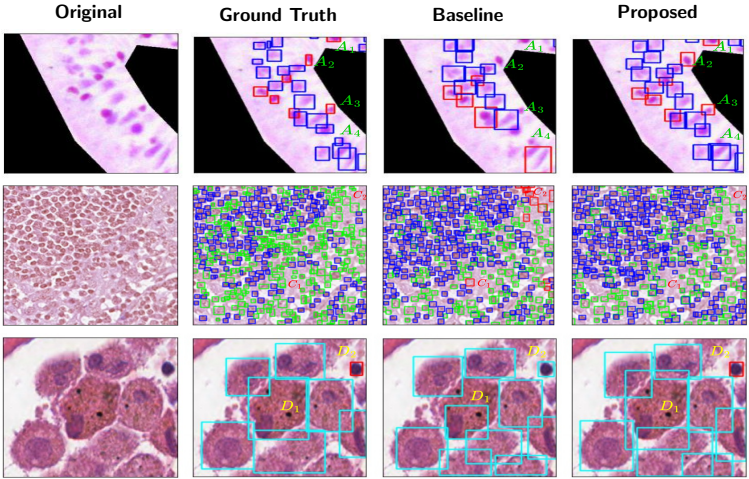


Figure: Qualitative performance of DeGPR.

Conclusion

- We address the multi-class cell detection and counting problem (MC2DC) in a limited data scenario for medical histopathological images.
- Deep Guided Posterior Regularization, introduces additional regularization terms to encourage the model to produce posterior distributions of features over predicted bounding boxes that resemble those of ground truth.
- DeGPR utilizes explicit features, provided by domain experts, and implicit features, learned through supervised contrastive loss on labeled data.
- We also contribute a new dataset consisting of 55 duodenum biopsies, which is valuable for predicting celiac disease.
- To facilitate further research, we have made our code and data publicly available.

Thanks

Code and Dataset

Github: https://github.com/dair-iitd/DeGPR

HuggingFace: https://huggingface.co/Aayushktyagi/DeGPR