Machine Learning Approach for Biopsy-Based Identification of Eosinophilic Esophagitis Reveals Importance of Global features

Tomer Czyzewski, Nati Daniel ⓘ, Mark Rochman, Julie M. Caldwell, Garrett A. Osswald, Margaret H. Collins, Marc E. Rothenberg, and Yonatan Savir ⓘ

Abstract—Goal: **Eosinophilic esophagitis (EoE) is an allergic inflammatory condition characterized by eosinophil accumulation in the esophageal mucosa. EoE diagnosis includes a manual assessment of eosinophil levels in mucosal biopsies–a time-consuming, laborious task that is difficult to standardize. One of the main challenges in automating this process, like many other biopsy-based diagnostics, is detecting features that are small relative to the size of the biopsy.** ***Results:*** **In this work, we utilized hematoxylin- and eosin-stained slides from esophageal biopsies from patients with active EoE and control subjects to develop a platform based on a deep convolutional neural network (DCNN) that can classify esophageal biopsies with an accuracy of 85%, sensitivity of 82.5%, and specificity of 87%. Moreover, by combining several downscaling and cropping strategies, we show that some of the features contributing to the correct classification are global rather than specific, local features.** ***Conclusions:*** **We report the ability of artificial intelligence to identify EoE using computer vision analysis of esophageal biopsy slides. Further, the DCNN features associated with EoE are based on not only local eosinophils but also global histologic changes. Our approach can be used for other conditions that rely on biopsy-based histologic diagnostics.**

Index Terms— **Decision support system, deep convolutional network, digital pathology, eosinophilic esophagitis, small features detection.**

Impact Statement— **Deep convolutional neural network, together with a systematic downscaling and cropping approach, can classify esophageal biopsies with high accuracy and reveals a global nature of the histologic features of eosinophilic esophagitis. Our approach of systematic analysis of the image size versus downscaling tradeoff can be used to improve disease classification performance and insight gathering in digital pathology.**

Tomer Czyzewski, Nati Daniel, and Yonatan Savir are with the Department of Physiology, Biophysics, and System Biology, Faculty of Medicine, Technion Israel Institute of Technology, Haifa 35254, Israel (e-mail: tomer.cz1.mail@gmail.com; nati.daniel@campus.technion.ac.il; yoni.savir@technion.ac.il).

Mark Rochman, Julie M. Caldwell, Garrett A. Osswald, and Marc E. Rothenberg are with the Division of Allergy, and Immunology, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH 45229 USA (e-mail: Mark.Rochman@cchmc.org; Julie.Caldwell1@cchmc.org; Garrett.Osswald@cchmc.org; Marc.Rothenberg@cchmc.org).

Margaret H. Collins is with the Division of Pathology, Cincinnati Children's Hospital Medical Center, Department of Pediatrics, University of Cincinnati College of Medicine, Cincinnati, OH 45229 USA (e-mail: Margaret.Collins@cchmc.org).

I. INTRODUCTION

Eosinophilic esophagitis (EoE) is a recently recognized chronic food allergic disease associated with esophageal specific inflammation characterized by high levels of eosinophils [1]. An allergic etiology is strongly supported by the efficacy of food elimination diets, the co-occurrence of EoE with other allergic diseases (e.g., asthma and atopic dermatitis), animal models demonstrating that experimental EoE can be induced by allergen exposure, and the necessity of allergic mediators of inflammation, such as Interleukin 5 and Interleukin 13, on the basis of animal models and clinical studies [1], [2]. Disease pathogenesis is driven by food hypersensitivity and allergic inflammation and multiple genetic and environmental factors [3]. Although a rare disease with a prevalence of approximately 1:2000 individuals, EoE is now the chief cause of chronic refractory dysphagia in adults and an emerging cause for vomiting, failure to thrive, and abdominal pain in children [1].

Histologically, EoE involves eosinophil-predominant inflammation of the esophageal mucosa. Microscopic examination of esophageal mucosal biopsies is a prerequisite for EoE diagnosis. During esophagogastroduodenoscopy (EGD), several esophageal biopsies are procured. These are then formalin-fixed, embedded, sectioned, and subjected to hematoxylin and eosin (H&E) staining [4], [5]. Subsequently, a pathologist examines the biopsies to determine the peak eosinophil count
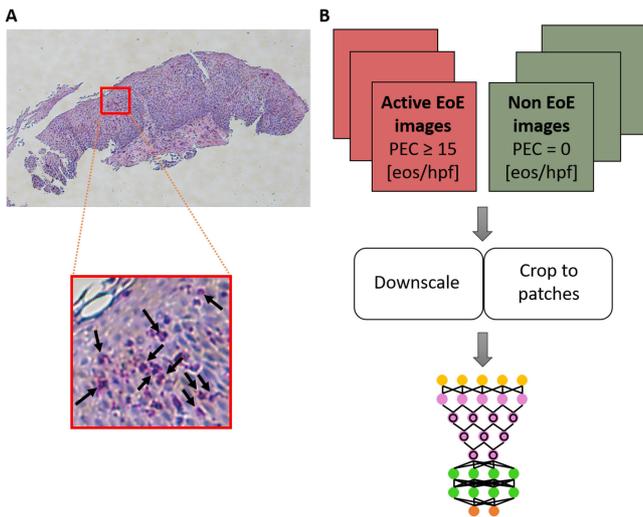
FIGURE 1. (A) Example of a full-size hematoxylin and eosin (H&E)-stained esophageal biopsy slide from a patient with active eosinophilic esophagitis (EoE). The red square marks an example of an area containing eosinophils (bright pink cells with purple nuclei; several examples are indicated by black arrows in the inset). (B) Schematics of the platform. Images (magnification $80\times$) of research slides (from one esophageal research biopsy per patient) are labeled as EoE or non-EoE on the basis of a pathologist's analysis of corresponding clinical slides associated with the same endoscopy during which the research biopsy was obtained. The full-size images are downscaled and/or cropped using various approaches to smaller images that are then used to train a deep convolutional neural network (DCNN). eos, eosinophils; hpf, high-power field; PEC, peak eosinophil count.

(PEC) [1], [2], [6] (Fig. 1). In addition to determining PEC, other histopathologic features of EoE include abnormalities of the structural cells, including epithelial cells and fibroblasts comprising the lamina propria. These features can be reliably assessed and quantified using the newly developed EoE Histology Scoring System (HSS) [7]. This system not only reports the presence or absence of the features but also takes into account grade (severity) and stage (extent). This scoring system is trainable across pathologists [7]. However, considerable disagreement can occur among certain observers, at least based on PEC [8], and even for trained observers, scoring esophageal biopsies requires a non-trivial time input.

During the last few years, deep learning and, in particular, deep convolutional neural networks (DCNNs) have become a significant component of computer vision. Unlike classical machine learning techniques, deep learning involves the net performing representation learning, which allows the machine to be fed raw data and to discover the representations needed for detection or classification automatically [9]–[12]. In particular, deep learning is used for the classification and diagnosis of conditions in which the diagnosis is based on histomorphology, such as cancer [12], [13]. However, the application of deep learning to medical applications poses two unique challenges: first, DCNN training requires a large number of images (hundreds to millions); and second, the size of the relevant objects within the images is small [14], [15].

Here, we developed a method based on DCNN and downscaling of esophageal biopsy images at different frequencies. By comparing the results of each frequency, we aimed to deduce whether the scattering is global (i.e., features appear diffusely throughout the tissue image) or local (i.e., features appear in only specific and/or discrete locations within the image). We developed a classifier that distinguishes between images of H&E-stained esophageal biopsies from patients with active EoE and non-EoE control patients with high accuracy. We show that some of the features that underlie the correct classification of disease are global in nature.

II. MATERIALS AND METHODS

A. Dateset

This study was performed under the Cincinnati Children's Hospital Medical Center (CCHMC) IRB protocol 2008-0090. Subjects undergoing endoscopy (EGD) for standard-of-care purposes agreed to donate additional gastrointestinal tissue biopsies for research purposes and to have their clinical, histologic, and demographic information stored in a private research database. One distal esophageal biopsy per patient was placed in 10% formalin; the tissue was then processed and embedded in paraffin. Sections ($4\mu m$) were mounted on glass slides and subjected to H&E staining, in a manner identical to the preparation of standard-of-care biopsies. Biopsies were viewed at $80\times$ magnification using the Olympus BX51 microscope, and one photograph of each biopsy was taken using the DP71 camera. Images were classified into categories on the basis of the clinical pathology report associated with the distal esophagus biopsies that were obtained for clinical analysis during the same endoscopy during which the biopsy for research purposes was procured. The clinical report is based on the observation of the pathologist that was available when the biopsy was taken. In the context of strictly counting eosinophils, the inter-observer and intra-observer correlation for reporting eosinophilic peak counts was reported to be more than 0.97 [16].

In this study, we used images defined as being derived from individuals with active EoE (biopsy with PEC ≥ 15 eosinophils [eos]/$400\times$ high-power field [hpf]) or from non-EoE control individuals (biopsy with PEC $= 0$ eos/hpf); (n $= 210$ non-EoE; n $= 210$ active EoE). The images were taken with digital microscopy at different sizes: 4140×3096 pixels, 2010×1548 pixels, or 1360×1024 pixels. In the original dataset, the number of images per category and at each size was not equal. Therefore, to avoid training bias, the images were randomly selected to build non-biased training and validation sets. In this new dataset, the number of images in each category was equal (training set: n $= 147$ active EoE, n $= 147$ non-EoE; validation set: n $= 63$ active EoE, n $= 63$ non-EoE). Additionally, the number of images per size was equal in each category (4140×3096: n $= 29$; 2010×1548: n $= 126$; 1360×1024: n $= 55$).

B. Downscale Approaches and Training

Two methods were employed to address the challenge of training on high-size images containing small features: first, downscaling the original image with the potential of losing the information associated with small features [14]; and second,

TABLE 1.

Whole Image Classification Results for Four Downscale and/or Crop Approaches. The Validation Cohort of Images (N = 63 Active EoE; N = 63 non-EoE) Was the Same for Each of the Classifiers. True Positive Rate (TPR; Number of Images Classified as Active EoE / Number of Active EoE Images X 100), True Negative Rate (TNR; Number of Images Classified as non-EoE / Number of non-EoE Images X 100), Accuracy (Number of Images Accurately Classified as Either Active EoE or non-EoE / Total Number of Images X 100), and Predicted Prevalence (Total Number of Images Classified as Active [I.E., True Positive + False Positive Number of Images] / Total Number of Images) for Each Method are Shown. DCNN, Deep Convolutional Neural Network. ACC, Accuracy

Original Image	Final DCNN input image size	Active EoE (TPR)	Non-EoE (TNR)	ACC	Predicted Prevalence (PP)
Full Image	1000x1000 (Downscale)	74.6%	96.8%	85.7%	0.39
Full Image	224x224 (Downscale)	65.1%	88.9%	77.0%	0.38
Patch = 448x448	224x224 (Downscale)	82.5%	87.3%	84.9%	0.48
Patch = 224x224	224x224	82.5%	77.8%	80.2%	0.52

FIGURE 2. Steps in processing esophageal biopsy images to produce patches. (A) A typical image of a hematoxylin and eosin (H&E) stained esophageal biopsy section obtained from an individual with active EoE. The image was taken at $80X$ magnification. (B) The same image after background removal with an illustration of tissue coverage criteria per patch size to meet the threshold for inclusion in training or validation sets. Box 1 (red): patch of 224×224 pixels with less than 10% tissue coverage. Box 2 (yellow): patch of 224×224 pixels with greater than 10% tissue coverage. Box 3 (red): patch of 448×448 pixels with less than 10% tissue coverage. Box 4 (yellow): patch of 448×448 pixels with greater than 10% tissue coverage.

dividing the images into smaller patches and analyzing each of the patches [17]. Although the second approach solves the image size challenge, if the relevant small feature (e.g., a local increase in eosinophil density) appears in only a few patches, many patches that do not contain the small feature are still labeled as positive. As a result, the false-positive prediction might significantly bias the final diagnosis. Yet, this method indicates whether the scatter of the features is global or local by carefully comparing it to a random classifier.

In this work, we used ResNet50, a residual network 50 layers deep [18]. Residual networks utilize skip-connections to transfer the output of a particular layer as input not only to the consecutive layer but also to subsequent layers. This property's main advantage is in coping with issues such as vanishing gradients and the degradation problem that are common when training very deep networks [19]. Four different DCNNs were trained, wherein each of the input image sizes was obtained differently: 1) cropping the full image to patches of 224×224 pixels (the optimal size for ResNet50), 2) cropping the full image to patches of 448×448 pixels and downscaling them to 224×224, 3) downscaling the original image to 224×224 pixels size, and 4) downscaling the original image to 1000×1000 pixels size (Table 1). This size was chosen because it represents nearly the maximum size possible for training on Nvidia 1080TI with a minimal mini-batch size of four images. Downscaling was done using bicubic interpolation.

Patches were cropped with a sliding window of the desired input (224×224, 448×448 pixels) with steps of half of the input size for overlay, covering the full original images (an example of a full image is shown in Fig. 2(A)). Subsequently, only patches that had more than 10% tissue comprising the patch were chosen for training and validation sets (Fig. 2(B)). All valid patches were used for training. Table S1 in the supplementary materials summarizes the number of images and patches for the validation and training sets for the various approaches. During training, rotation, translation, and reflection augmentations were
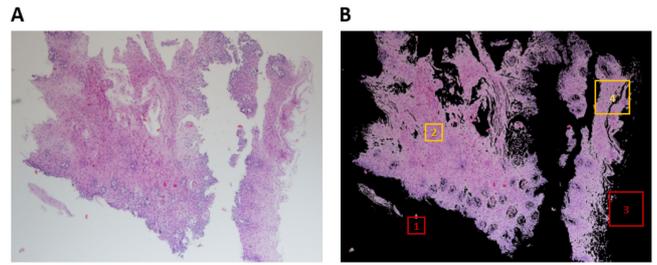
performed. We used imageDataAugmenter, a MATLAB-based method, to augment the training set images. Each image in the training set was duplicated and underwent a combination of rotation, translation, and reflection. Rotation angles were random in the range [0 90] degrees, reflection was either horizontally or vertically with 50% probability, and x and y translation were drawn randomly in the range [0 100] pixels (image size of 1000 $\times$ 1000) or [0 20] (image size of 448×448 or 224×224).

III. Results

Table 1 summarizes the whole image classification results for the four downscale and/or crop approaches employed. First, we downscaled the original images to two different input image sizes. If the majority of the information that defines the condition were local, we would expect that downscaling, resulting in smooth local features, would have a significant effect on the classification quality. Surprisingly, we found that downscaling the original images to a size of 1000×1000 did not result in a random classification, but instead resulted in a true positive rate (TPR) of 74.6% and a true negative rate (TNR) of 96.8%. These results suggest that some of the information that defines the condition is local but is large enough to sustain the downscaling; alternatively, the information could be global. The bias towards negative classification (predicted prevalence [PP] <0.5), as indicated by the PP of 0.39, suggests that the information that determines the condition is more local, leading to more positive-labeled images having the same feature as negative-labeled images. Downscaling the full images even further to a size of 224×224 reduced both the TPR and the TNR. Yet, consistent with the hypothesis that the information that defines the positive images is more sensitive to downscaling, the PP remained similar, and the TPR was reduced more than the TNR ($\Delta 9.5\%$ and $\Delta 7.9\%$, respectively). It is insightful to examine quantitively the effect of the downscaling factor (the ratio between the original image area and input images area) on accuracy (Supplementary Materials, section II, Fig S1A). As

the downscaling factor is more than ~5, there is a decrease in accuracy.

Next, we classified the whole images according to the sub-classification of their patches. The predicted label assigned to the whole image (i.e., active EoE or non-EoE) resulted from the majority vote of the predicted labels of its patches (i.e., if $\geq 50\%$ of patches were classified as active EoE, the whole image was classified as active EoE; if $\geq 50\%$ of patches were classified as non-EoE, the whole image was classified as non-EoE). First, each image was parsed into patches, each with a size of 448×448 that were then each downscaled to a size of 224×224. In this case, no substantial classification bias resulted; the PP of 0.48 and the TPR of 82.5% increased substantially compared to the two downscaling methods described previously (Table 1).

Using patches of 224×224 that did not undergo downscaling yielded a similar TPR of 82.5%; however, the TNR decreased to 77.8%. Breaking down the effect of image size, which is proportional to the number of patches, reveals a monotonically increasing relation between the image size and accuracy. Yet, cropping into 448×448 patches and downscale to 224×224 gives better or equal accuracy, for all image sizes, than cropping into 224×224 patches (Supplementary Materials, section II, Fig. S1B). This is likely due to the inherent tradeoff between the local and global information contained within the images. If an image is larger, it contains more global information, but the downscaling that is required prior to its input into the net is larger; thus, small features are smoothed out to a greater degree. In our case, using a 448×448 patch with downscaling provided a better TNR of 87.3% than did using smaller patches of 224×224 without downscaling.

Figure 3 summarizes the effect of the initial patch size and downscaling factor in the receiver operating characteristic (ROC) space. Applying other aggregation methods, such as hierarchal clustering or different voting thresholds, resulted in similar results (Supplementary Materials, section III, Fig. S2). We also benchmarked the DCNN classification results using standard well-known 20 textural features [20], [21], and three baseline classification methods: linear discriminant analysis, logistic regression, and linear SVM (Supplementary Materials, section IV). The DCNN performs significantly better than these standard classification approaches (Supplementary Materials, section IV, Table S3). Additional performance measures such as recall, precision, and F1-score are summarized in Table S4 in the supplementary materials section V.

To further analyze the tradeoff between locality and down-scale factor, we evaluated the classification performance of the patches themselves (Table 2). The results are consistent with the whole image majority vote classification. In particular, both the TNR of 79.7% and TPR of 77.0% of the 448×448 patch downscaled to 224×224 are higher than those of the non-scaled 224×224 patch. These results indicate that incorporating more information in a patch is more important than downscaling by a factor of two and supports the notion that global information drives the classification for EoE.

To determine the effect of locality on the classification, we compared the distribution of prediction probability for patches with a size of 224×224 that did not undergo downscaling in
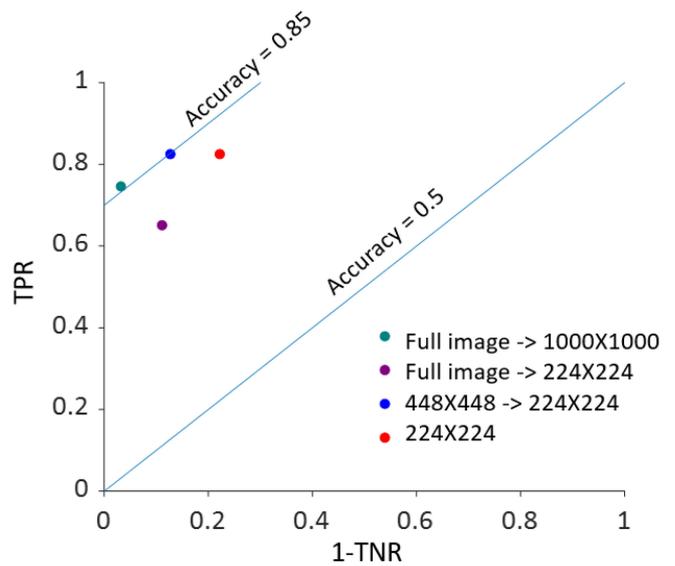

FIGURE 3. Classification results as a function of initial image size and downscaling factor in the receiver operating characteristic (ROC) space. For each of the four downscale and/or crop approaches utilized to analyze the validation cohort of images (n = 63 active EoE; n = 63 non-EoE), the true positive rate (TPR) vs. (1 - the true negative rate [TNR]) with TPR and TNR expressed as proportions is graphed. Blue lines highlight accuracy measurements of 50% and 85% expressed as proportions.

TABLE 2.
CLASSIFICATION RESULTS FOR INDIVIDUAL PATCHES. THE VALIDATION COHORT OF IMAGES (N = 63 ACTIVE EoE; N = 63 NON-EoE) WAS SUBJECTED TO CROPPING INTO PATCHES WITH THE INDICATED PIXEL SIZES AND DOWNSCALED WHEN INDICATED. TRUE POSITIVE RATE (TPR; NUMBER OF PATCHES CLASSIFIED AS ACTIVE EoE / NUMBER OF ACTIVE EoE PATCHES X 100), TRUE NEGATIVE RATE (TNR; NUMBER OF PATCHES CLASSIFIED AS NON-EoE / NUMBER OF NON-EoE PATCHES X 100), ACCURACY (NUMBER OF PATCHES ACCURATELY CLASSIFIED AS EITHER ACTIVE EoE OR NON-EoE / TOTAL NUMBER OF PATCHES X 100), AND PREDICTED PREVALENCE (TOTAL NUMBER OF IMAGES CLASSIFIED AS ACTIVE [I.E., TRUE POSITIVE + FALSE POSITIVE NUMBER OF IMAGES] / TOTAL NUMBER OF IMAGES) FOR EACH PATCH SIZE AND DOWNSCALING METHOD (IF APPLICABLE) ARE SHOWN. DCNN, DEEP CONVOLUTIONAL NEURAL NETWORK; TPR, TRUE POSITIVE RATE; TNR, TRUE NEGATIVE RATE. ACC, ACCURACY

Original Image	Final DCNN input image size	Active EoE (TPR)	Non-EoE (TNR)	ACC	Predicted Prevalence (PP)
Patch = 448x448	224x224 (Downscale)	77.0%	79.7%	78.3%	0.49
Patch = 224x224	224x224	73.3%	75.2%	74.2%	0.49

two cases. In the first, each patch was labeled with the same label as the original image from which it was derived. In the second, each patch was assigned a random label.

Figure 4 shows the distribution for each case. In the case in which the patch labels are true (Fig. 4(A), (B)), the distribution is bi-modal. In the case in which the patch labels are random (Fig. 4(C), (D)), most of the patches are ambiguous, and thus the distribution is unimodal around 0.5. These collective case findings suggest that most of the patches that are classified correctly are not ambiguous. This indicates that the local patch
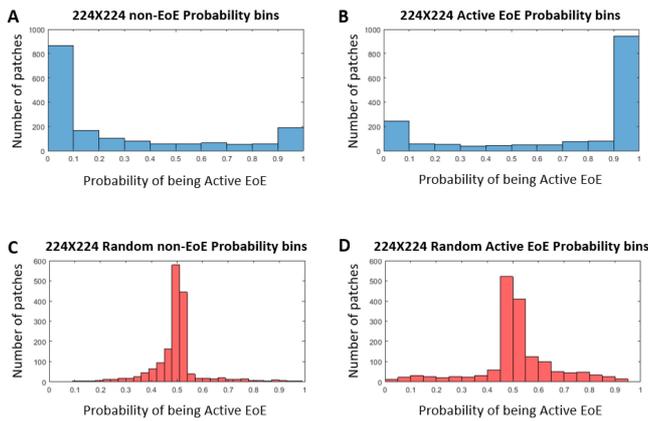
FIGURE 4. Prediction ability of nonrandom (blue) and random (red) classifier. (A) 224 × 224: histogram of the number of patches derived from non-EoE images vs. the probability that they will be classified as active EoE by the nonrandom classifier. (B) 224 × 224: histogram of the number of patches derived from active EoE images vs. the probability that they will be classified as active EoE by the nonrandom classifier. (C) Random 224 × 224: histogram of the number of patches derived from non-EoE-labeled images vs. the probability that they will be classified as active EoE by the random classifier. (D) Random 224 × 224: histogram of the number of patches derived from active EoE-labeled images vs. the probability that they will be classified as active EoE by the random classifier.

FIGURE 5. Schematic of various potential distributions of local patterns within an esophageal biopsy section. An esophageal biopsy image is shown; red ovals denote a local feature that contributes to disease diagnosis. (A) Local pattern confined to a specific place in the tissue. (B) Local pattern distributed at the edge of the tissue. (C) Local pattern restricted to only half of the tissue. (D) Global pattern spread all over the tissue.

labeling carries information that is relevant for the majority of the patches.

IV. DISCUSSION AND CONCLUSION

One of the main challenges in digital pathology is that the features of the conditions are very small compared with the size of the sample. This feature-sample size disparity leads to an inherent tradeoff between the size of the analyzed image and the downscaling factor. In the case of small, local features, visualizing the image as smaller patches may impede the classification because most of the patches will not include the small, local features. However, if local features are the primary source of information about the condition, downscaling the whole image may smooth them out.

Herein, we used DCNN and different downscaling and/or cropping approaches to achieve ∼85% accuracy in distinguishing active EoE from non-EoE esophageal biopsies, despite the relatively small number of labeled images utilized for training (n = 147 active EoE and n = 147 non-EoE). Although labeling relied primarily on a local feature (PEC ≥ 15 eos/hpf), our results support that EoE is also associated with additional global histopathologic features that are learned by the classifier. Figure 5 illustrates possible scatter patterns for features that contribute to disease diagnosis. Of note, the features could be clustered locally (e.g., a local increase in density of eosinophils), or they could be distributed uniformly throughout the tissue (e.g., morphology of structural cells comprising the tissue).

The fact that images that were cropped into patches but were downscaled by a factor of greater than 10 (in terms of the number of pixels) provided low TPR, suggests that the features associated with the condition were not big enough for the classification task. Howe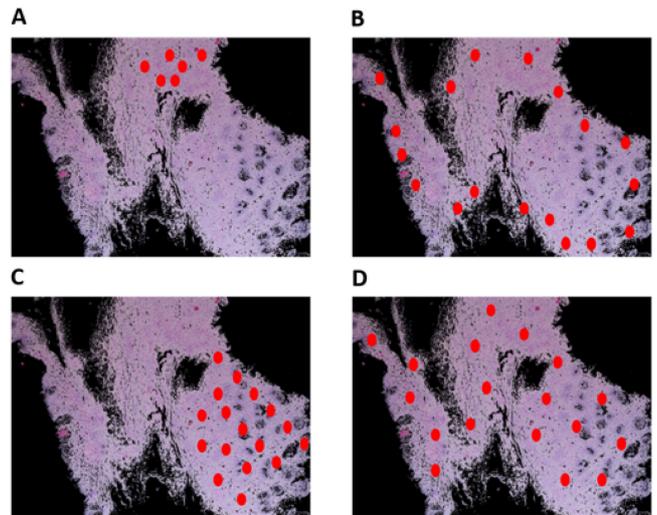ver, if the features were distributed only locally (e.g., Fig. 5(A)–(C)), many patches cropped from the whole image would not include the features, and thus the classification according to patches would fail. However, in this study of EoE, most of these cropped patches were labeled correctly. Moreover, the classification was better with 448 × 448 patches downscaled to 224 × 224 than non-scaled 224 × 224 patches, suggesting presence of global features (Fig. 5(D)). Our results thus indicate that although the original labeling was based primarily on local features, additional global features are associated with EoE (Fig. 5(D)). This global information allows a classification with minimal PP bias (PP 0.49) and with only a small number of images.

In this work, we used an approach in which the label is global - the entire slide is labeled according to the patient condition (e.g., whether a patient is active or not) - and the network is trained without local labeling. To improve the accuracy, semantic information can be incorporated to estimate the number of eosinophils directly [22]. Another approach is to score the images not only according to their PEC, but also to account for additional features such as basal zone hyperplasia and dilated intercellular spaces [7].

Our work highlights the importance of systematic analysis of the image size vs. downscaling tradeoff, particularly in digital pathology, for improving classification and gaining insight into the features' spatial distribution underlying a condition. These findings present an initial artificial intelligence approach to diagnosing EoE using digital microscopy and have implications for analyzing other biopsy-based disease diagnoses.

IV. ACKNOWLEDGMENT

REFERENCES

[1] E. S. Dellon *et al.*, "Updated international consensus diagnostic criteria for eosinophilic esophagitis," *Gastroenterology*, vol. 155, pp. 1022–1033, Oct. 2018.

[2] J. D. Sherrill and M. E. Rothenberg, "Genetic dissection of eosinophilic esophagitis provides insight into disease pathogenesis and treatment strategies," *J. Allergy Clin. Immunol.,* vol. 128, no. 1, pp. 23–32, Jul. 2011.

[3] K. M. O'Shea *et al.*, "Pathophysiology of eosinophilic esophagitis," *Gastroenterology*, vol. 154, pp. 333–345, Jan. 2018.

[4] J. D. Bancroft and M. Gamble, *Theory and Practice of Histological Techniques*. Amsterdam, The Netherlands: Elsevier Health Sciences, 2008.

[5] G. Gill, *Cytopreparation: Principles & Practice*. Berlin, Germany: Springer Science & Business Media, 2012.

[6] E. S. Dellon, K. J. Fritchie, T. C. Rubinas, J. T. Woosley, and N. J. Shaheen, "Inter- and intraobserver reliability and validation of a new method for determination of eosinophil counts in patients with esophageal eosinophilia," *Dig. Dis. Sci.*, vol. 55, pp. 1940–1949, Jul. 2010.

[7] M. H. Collins *et al.*, "Newly developed and validated eosinophilic esophagitis histology scoring system and evidence that it outperforms peak eosinophil count for disease diagnosis and monitoring," *Dis. Esophagus*, vol. 30, pp. 1–8, Mar. 2017.

[8] E. M. Stucke, K. E. Clarridge, M. H. Collins, C. J. Henderson, L. J. Martin, and M. E. Rothenberg, "Value of an additional review for eosinophil quantification in esophageal biopsies," *J. Pediatr. Gastroenterol. Nutr.*, vol. 61, pp. 65–68, Jul. 2015.

[9] M. Manoj krishna, M. Neelima, A. M. Harshali, and M. Venu GopalaRao, "Image classification using deep learning," *Int. J. Eng. Technol.*, vol. 7, pp. 614–617, Mar. 2018.

[10] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, 2017.

[11] R. Fakoor, A. Nazi, and M. Huber, "Using deep learning to enhance cancer diagnosis and classification," in *Proc. 30th Int. Conf. Mach. Learn.*, 2013.

[12] J. Wang, J. D. MacKenzie, R. Ramachandran, and D. Z. Chen, "A deep learning approach for semantic segmentation in histology tissue images," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.,* S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, Eds., *Lecture Notes in Computer Science*, Cham, Switzerland: Springer International Publishing, 2016, pp. 176–184.

[13] U. Djuric, G. Zadeh, K. Aldape, and P. Diamandis, "Precision histology: How deep learning is poised to revitalize histomorphology for personalized cancer care,*" NPJ Precis. Oncol.*, vol. 1, p. 22, Dec. 2017.

[14] K. J. Geras, S. Wolfson, Y. Shen, S. G. Kim, L. Moy, and K. Cho, "High-resolution breast cancer screening with multi-view deep convolutional neural networks," *Comput. Res. Repository*, vol. 1703, 2018, Art. no. 07047.

[15] A. Shawahna, S. M. Sait, and A. El-Maleh, "FPGA-based accelerators of deep learning networks for learning and classification: A review," *IEEE Access*, vol. 7, pp. 7823–7859, 2019.

[16] A. Vanstapel, T. Vanuytsel, and G. D. Hertogh, "Eosinophilic peak counts in eosinophilic esophagitis: A retrospective study," *Acta Gastro-Enterologica Belgica*, vol. 82, no.2, pp. 243–250, 2019.

[17] V. Kovalev, A. Kalinovsky, and V. Liauchuk, "Deep learning in big image data: Histology image classification for breast cancer diagnosis protein docking by deep neural networks view project UAV: Back to base problem view project," in *Proc. Int. Conf. Big Data Adv. Anal.*, 2016, pp. 15–17.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016, 2016, pp. 770–778.

[19] Z. Wu, C. Shen, and A. van denHengel, "Wider or deeper: Revisiting the RESNET model for visual recognition," *Pattern Recognit.*, vol. 90, pp. 119–133, 2019.

[20] R. Haralick and L. Shapiro, *Computer and Robot Vision.* Reading, MA, USA: Addison-Wesley, 1st ed., 1992.

[21] R. M. Haralick, I. Dinstein, and K. Shanmugam, "Textural features for image classification," *IEEE Trans. Syst., Man Cybern.*, vol. SMC- 3, no. 6, pp. 610–621, Nov. 1973.

[22] N. Daniel, A. Larey, E. Aknin, G. A. Osswald, and J. Caldwell *et al.*, "PECNet : A deep multi-label segmentation network for eosinophilic esophagitis biopsy diagnostics," 2021, *arXiv:2103.02015*.