

Data Descriptor

Not peer-reviewed version

---

# Dataset and AI Workflow for Deep Learning Image Classification of Ulcerative Colitis and Colorectal Cancer

---

[Joaquim Carreras](#)\*, [Giovanna Roncador](#), [Rifat Hamoudi](#)

Posted Date: 15 April 2025

doi: 10.20944/preprints202412.1201.v4

Keywords: ulcerative colitis; colorectal cancer; adenocarcinoma; artificial intelligence; deep learning; convolutional neural network; transfer learning; TOX2; LAIR1; H&E; whole-slide images; ResNet-18



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

*Data Descriptor*

# Dataset and AI Workflow for Deep Learning Image Classification of Ulcerative Colitis and Colorectal Cancer

Joaquim Carreras <sup>1,\*</sup>, Giovanna Roncador <sup>2</sup> and Rifat Hamoudi <sup>3,4,5,6,7</sup>

<sup>1</sup> Department of Pathology, Tokai University, School of Medicine; 143 Shimokasuya, Isehara, Kanagawa, 259-1193 Japan; joaquim.carreras@tokai.ac.jp; joaquim.carrerasesteban@cantab.net

<sup>2</sup> Monoclonal Antibodies Unit, Spanish National Cancer Research Center (CNIO), Melchor Fernandez Almagro 3, 28029 Madrid, Spain; groncador@cnio.es

<sup>3</sup> Department of Clinical Sciences, College of Medicine, University of Sharjah, Sharjah P.O. Box 27272, United Arab Emirates; rhamoudi@sharjah.ac.ae

<sup>4</sup> Biomedically Informed Artificial Intelligence Laboratory (BIMAI-Lab), University of Sharjah, Sharjah P.O. Box 27272, United Arab Emirates

<sup>5</sup> Center of Excellence for Precision Medicine, University of Sharjah, Sharjah P.O. Box 27272, United Arab Emirates

<sup>6</sup> Division of Surgery and Interventional Science, University College London, London NW3 2PF, UK

<sup>7</sup> ASPIRE Precision Medicine Research Institute Abu Dhabi, University of Sharjah, Sharjah P.O. Box 27272, United Arab Emirates

\* Correspondence: joaquim.carreras@tokai.ac.jp; Tel.: +81-463-93-1121; Fax.: +81-463-91-1370

**Abstract:** Inflammatory bowel disease (IBD) is a chronic inflammatory condition of the gastrointestinal tract characterized by the deregulation of immuno-oncology markers. IBD includes ulcerative colitis and Chron disease. Chronic active inflammation is a risk factor for the development of colorectal cancer (CRC). Deep learning is a form of machine learning that is applicable to computer vision, and it includes algorithms and workflows used for image processing, analysis, visualization, and algorithm development. This technical note of data descriptor and methods type describes a dataset of histological images of ulcerative colitis, CRC (adenocarcinoma), and colon control. The samples were stained with hematoxylin and eosin (H&E), and immunohistochemically analyzed for LAIR1 and TOX2 markers. The methods used for collecting, processing, and analyzing scientific data using convolutional neural networks (CNNs), where the dataset can be found, and information about its use is also described. This note is a companion to the DOI: 10.3390/cancers16244230.

**Keywords:** ulcerative colitis; colorectal cancer; adenocarcinoma; artificial intelligence; deep learning; convolutional neural network; transfer learning; TOX2; LAIR1; H&E; whole-slide images; ResNet-18

## 1. Summary

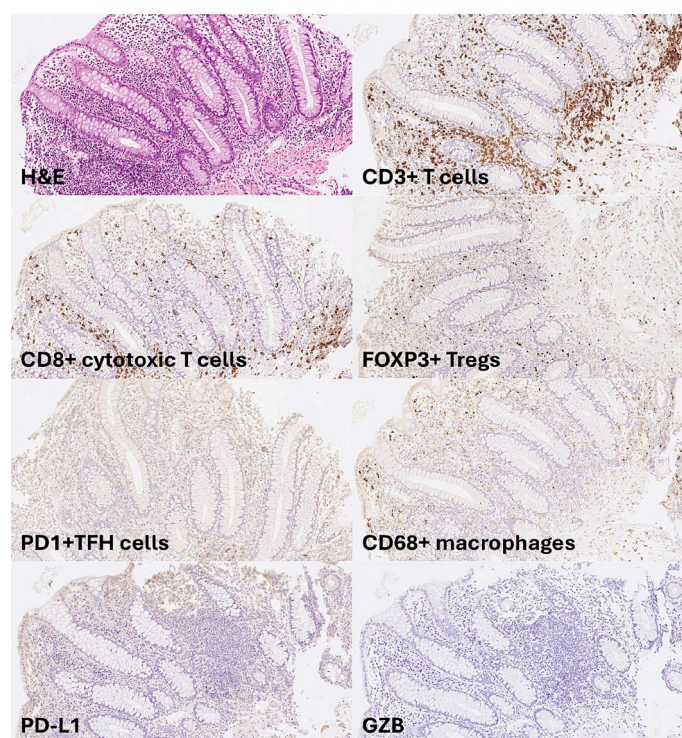
### 1.1. Background of Ulcerative Colitis

Inflammatory bowel disease (IBD) is a chronic inflammatory condition of the gastrointestinal tract with systemic repercussions that is characterized by relapsing and remitting episodes of inflammation [1,2]. IBD includes 2 types: ulcerative colitis, which affects the colon, and Chron disease, which can affect any part of the gastrointestinal tract from the mouth until the perianal area [1–9].

The causes and pathogenesis of IBD are unclear. However, it appears to be a combination of factors from the environment, including the microbiome [10,11], genetic susceptibility [12,13], and immune system (both systemic and gut) [13–16]. The prevalence of IBD has increased globally in recent decades, particularly in industrialized countries [17–21]. Clinical risk factors include smoking

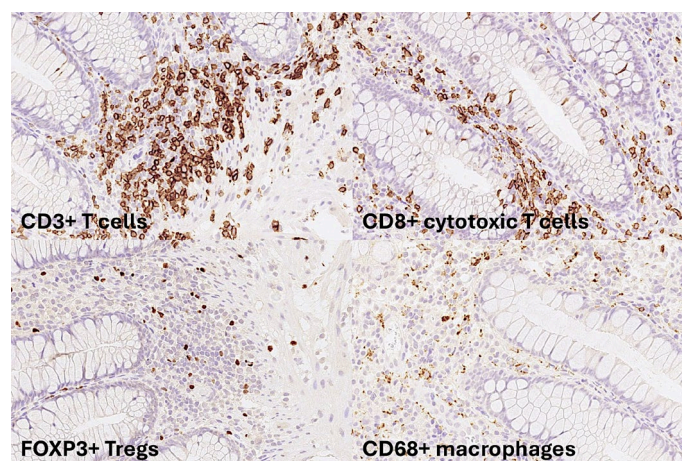
[22], low physical activity, low fiber intake, high fats, and low vitamin D [23], sleep deprivation, previous acute gastroenteritis, antibiotic use, and early life exposures [24].

In normal and healthy conditions, the intestinal barrier is maintained by the mucus layer and epithelial cells, which create bonds using tight junctions. This barrier is supported by the presence of IgA and several antimicrobial factors. The immune response is initiated by dendritic cells, which acquire, process, and present antigens to B and T lymphocytes [25]. The cause of ulcerative colitis is still not completely understood. Ulcerative colitis is associated with mucosal barrier damage, microflora alteration, and an uncontrolled immune reaction. Several subsets of CD4-positive T cells are believed to participate in the pathogenesis of UC, including Th1, Th2, Th9, Th17, Th22, TFh lymphocytes, and Tregs [26,27]. Th9 are associated with enterocyte apoptosis and inhibit mucosal healing [25,28]. IL-13 and NK/T cells are also involved in epithelial injury [25]. Innate immune cells contribute to cytokine production and continuous inflammation [14,15,29]. Injury of the mucosa is associated with dysbiosis [30], which is defined by an alteration of the composition of the mucosa, with changes of the diversity, increased potentially pathogenic bacteria, and reduced beneficial bacteria [31]. Figures 1 and 2 show examples of ulcerative colitis and IHC of the cells of the microenvironment involved in the pathogenesis.



**Figure 1.** Pathogenesis of ulcerative colitis and cells of the immune microenvironment. Ulcerative colitis is a chronic inflammatory condition affecting the colon. The incidence of ulcerative colitis has been increasing in the recent decades. The pathogenesis is multifactorial, including genetic predisposition, dysregulation of the immune tolerance, and mucosal homeostasis. Several immune cells are involved. This figure shows the immunohistochemical staining of CD3-positive T lymphocytes (cells), CD8-positive cytotoxic T lymphocytes, FOXP3-positive regulatory T lymphocytes (Tregs), PD-1-positive follicular T helper cells (TFH), and CD68-positive macrophages. H&E, hematoxylin and eosin staining.





**Figure 2.** Detailed visualization of some immune microenvironment components of ulcerative colitis. This figure shows the details of the tissue distribution of some of the immune microenvironment cells in ulcerative colitis at higher magnification. The lamina propria is characterized variable infiltration of CD3+T lymphocytes and CD68+macrophages. Within the CD3+T lymphocytes, cytotoxic CD8+T lymphocytes and FOXP3+regulatory T lymphocytes (Tregs) are easily found. .

The disease activity of ulcerative colitis is being recorded to properly treat patients and perform clinical trials. Usually, the severity is classified as mild, moderate, or severe. The available classifications include the Mayo [32], Montreal [33], Truelove, and Witts classifications [34]. The grade of mucosal inflammation can be assessed using endoscopy to record the degree of mucosal ulceration and the extent of disease. This can be graded using the Mayo score [32] and Baron score. The degree of histological severity can be evaluated using the Geboes score [35].

Microscopic (histologic) evaluation of ulcerative colitis shows active chronic colitis in patients with untreated disease. Disease activity is defined by neutrophil infiltration of the lamina propria, cryptitis, crypt abscess, and ulceration. The inflammation is limited to the mucosa and submucosa. Compared with Crohn disease, this condition is associated with a lack of transmural inflammation, granuloma, and fissuring ulcers. Dysplasia may be present in patients with long-term disease [36–38].

Disease activity can be evaluated using various parameters. Inactive disease lacks neutrophils. If the activity affects less than 50% of the mucosa, it is called mild. When it is >50% and there are crypt abscesses, it is called moderate. Finally, severe activity is characterized by surface ulcerations and erosion [39].

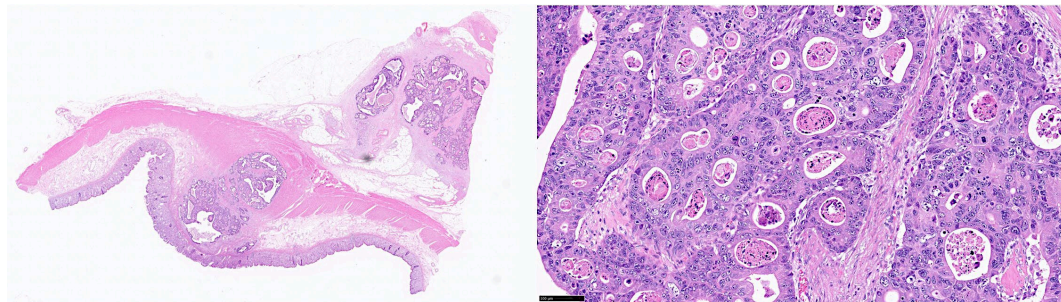
The first-line therapy for the induction and maintenance of remission of mild to moderate UC is 5-aminosalicylic acid [40]. The initial therapy for ulcerative colitis is topical mesalamine, which is available as a suppository or enema, and it is recommended for ulcerative proctitis or proctosigmoiditis. If topical mesalamine is not tolerated, an alternative therapy is topical glucocorticoids (i.e. hydrocortisone). A combination of 5-ASA and rectal mesalamine may be used for left-sided or extensive colitis. The 5-ASA formulations include mesalamine, sulfasalazine, and diazo-bonded 5-ASA. Patients who do not respond may subsequently be treated with glucocorticoids such as budesonide, which may escalate to prednisone or biological agents (anti-TNF, anti-integrin, anti-IL12/23, S1P modulators, JAK inhibitors) [40–43]. Patients who do not respond to oral glucocorticoids may require intravenous glucocorticoids. In some cases, surgery is performed [40–43]. Fecal microbiota transplantation is also available [44].

## 1.2. Background of Colorectal Cancer

Colorectal cancer (CRC) is a common disease characterized by environmental and genetic factors. It is the third most frequently diagnosed cancer worldwide [45]. There are several factors associated with early onset of CRC, including hereditary syndromes (familial adenomatous polyposis and



Lynch syndrome), metabolic dysregulation, history of CRC in first-degree relative; alcohol consumption, lack of regular use of nonsteroidal anti-inflammatory drugs (NSAIDs), and vitamin D intake. Other known risk factors of CRC include inflammatory bowel disease (ulcerative colitis and Crohn disease), abdominopelvic radiation, obesity, diabetes mellitus, insulin resistance, red and processed meat, and tobacco [46–49]. The diagnosis is usually made by colonoscopy, and treatment includes surgical resection, adjuvant chemotherapy, and postoperative radiation therapy [46]. Adenocarcinoma is the most common histological subtype of CRC. Adenocarcinoma is a glandular neoplasm. Most CRC cases are moderately differentiated with simple, complex, or slightly irregular tubules and loss of nuclear polarity. The glands are often cribriform, with necrosis, inflammatory cells, and marked desmoplasia [50–52] (Figure 3).



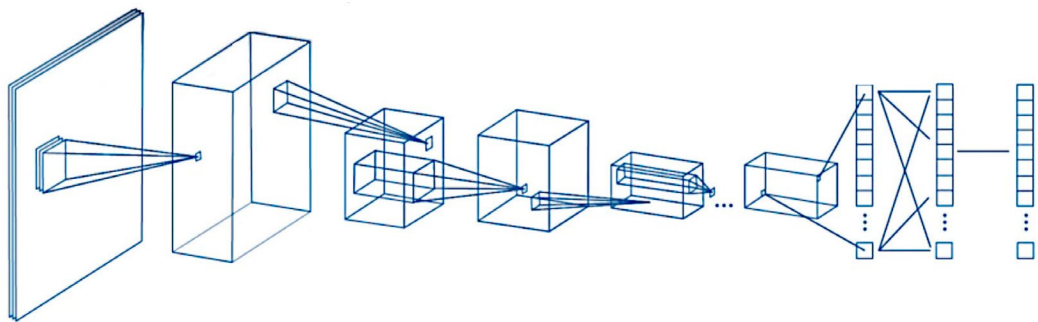
**Figure 3.** Histological features of colorectal cancer. The most common histological subtype of colorectal cancer is adenocarcinoma.

### 1.3. Background of Computer Vision for Deep Learning Image Classification

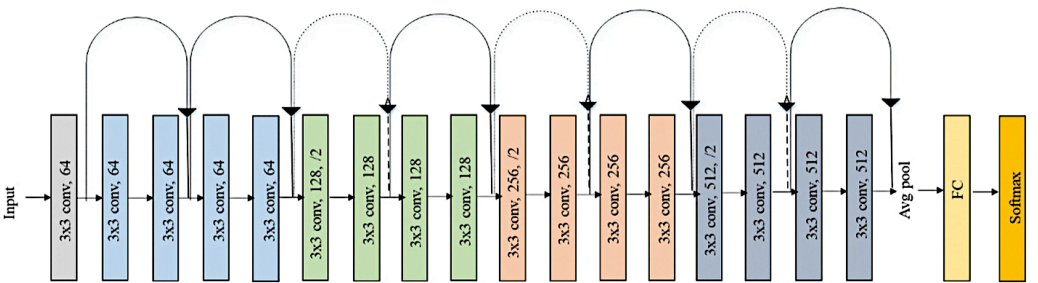
Image processing involves the algorithms and workflows of image processing, analysis, visualization, and algorithm development. Computer vision workflows [53] with deep learning include several types of functions, such as image classification, object detection, and instance segmentation, automated visual inspection, semantic segmentation, and video classification [54–56].

A pretrained neural network that has already learned how to extract the characteristics of natural images can be used as a starting point to handle new images. Using a pretrained image classification network [57], the learning time is shorter, and training is usually easier. The transfer learning process takes layers from an already trained neural network and calibrates the parameters on a new dataset [54–56,58–65]. The analysis includes a series of steps, including preprocessing the data, import pretrained networks from platforms such as TensorFlow 2 [66], TensorFlow-Keras [67], Pythorch [68], and ONNX [69], building the network, selecting training options, improving the network performance by tuning hyperparameters, visualizing and verifying network behavior during and after training, and exporting the network to other platforms if necessary[54–56,58–65].

The most important features of a neural network are accuracy, speed, and size. A good neural network is characterized by being fast and having good performance (i.e. accurate). Neural networks that are accurate in ImageNet can be used to classify other images using transfer learning or feature extraction. The best examples of neural networks, the best examples are GoogLeNet, ResNet-18, NobileNet, ResNet-50, ResNet-101, and Inception-v3 [56]. Figure 4 shows the general design of the CNN, and Figure 5 shows the original RestNet-18 architecture [70].



**Figure 4.** Structure of a convolutional neural network (CNN). The CNN algorithm is characterized by taking an input image and assigning weights and biases to different components. Then, the deep learning algorithm performs image classification. The CNN comprises of 3 main layers: the convolutional, pooling, and fully connected layers.



**Figure 5.** Original ResNet-18 architecture.

1.4. Dataset and Research Project Description

Ulcerative colitis is a chronic inflammatory bowel disease associated with a high risk of colorectal cancer. This study used convolutional neural networks and computer vision to classify histological images of ulcerative colitis, colorectal cancer (adenocarcinoma), and colon control.

The series included 35 patients with ulcerative colitis, 18 with colorectal cancer, and 21 with colon control. Hematoxylin and eosin (H&E) glass slides were converted into high-resolution digital data by high-speed scanning. The whole-tissue slides were split into image patches of 224×224 pixels at 200× magnification and 150 dpi, and the ResNet-18 network was retrained to classify the 3 types of diagnosis. This transfer learning experiment also used other pretrained CNNs for performance comparison, and the gradient-weighted class activation mapping (Grad-CAM) heatmap technique was used to understand the classification decisions.

Additionally, immunohistochemical analysis of 2 new immuno-oncology markers, LAIR1 and TOX2 were analyzed in ulcerative colitis to differentiate between mesalazine-responsive and steroid-requiring patients. LAIR1 an inhibitory receptor that plays a constitutive negative regulatory role in the cytolytic function of natural killer (NK) cells, B lymphocytes and T lymphocytes [71–73]. TOX2 is a new marker similar to PD-1 in the coinhibitory pathway [71,74].

Statistical analyses showed that steroid-requiring ulcerative colitis was characterized by higher endoscopic Baron and histologic Geboes scores and LAIR1 expression, and lower TOX2 expression in isolated lymphoid follicles. The CNN managed to classify the 3 diagnoses with >99% accuracy, and the Grad-CAM confirmed which parts of the images were relevant for the classification. In conclusion, the study proposed that CNNs are essential tools for deep learning image recognition.

These results were published as preprints [75].

This paper is a companion manuscript of the recently published article “Ulcerative Colitis, LAIR1 and TOX2 Expression, and Colorectal Cancer Deep Learning Image Classification Using

Convolutional Neural Networks” published in Cancers 2024, 16, 4230. <https://doi.org/10.3390/cancers16244230>.

## 2. Data Description

The dataset contains image patches of ulcerative colitis, colorectal cancer, and colon control. The image patches were split from whole-tissue images and had a size of  $224 \times 224 \times 3$ . The original magnification of the images was  $200\times$  and a dpi of 150. The image patches are anonymized. After splitting the images, the patches were filtered. The criteria were as follows: (1) image patches of only  $243 \times 243$  size; (2) image patches of more than 5-31 KB that contain tissue at least 20-30% of viable tissue; (3) image patches with diagnostic areas; (4) image patches without artifacts, including broken tissue, folded areas, incorrectly stained tissue, and smashed/crushed tissue. Steps 3 and 4 were manually curated by a pathology specialist (MD PhD).

The investigations were carried out according to the Declaration of Helsinki (website: <https://www.wma.net/policies-post/wma-declaration-of-helsinki/>; last accessed on December 12, 2024). Approval from an ethics committee was obtained before undertaking the research (protocol code IRB14R-080, IRB20-156, and 13R-119).

## 3. Methods

### 3.1. Formalin Tissue Fixation and Paraffin Embedding

Samples were fixed and embedded for long-term storage and immunostaining using fixatives to preserve tissues. The protocol included neutral buffered formalin, 80% alcohol, 95% alcohol, 100% alcohol, xylene, and paraffin pellets. Processing was performed using a HistoCore Peloris 3 premium tissue processing system (#13B2X10268PELOR3, Leica Biosystems K.K., Tokyo, Japan) (Figure 6).



**Figure 6.** Tissue processor. Processing was performed using the HistoCore Peloris 3 tissue processing system.

The processed tissue was embedded in paraffin, forming paraffin blocs ready for sectioning using a tissue embedder (Tissue-Tek® TEC™ 6 Embedding Console System, Sakura Finetek Japan Co., Ltd., Tokyo, Japan) (Figure 7).





**Figure 7.** Tissue embedder. The processed tissue was embedded in paraffin, forming paraffin blocs for sectioning using a tissue embedder (Tissue-Tek® TEC™ 6 Embedding Console System).

### 3.2. Sectioning of Paraffin-Embedded Tissue

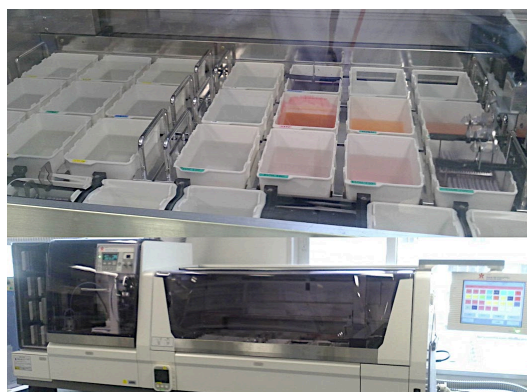
Paraffin blocks were trimmed at  $2\mu\text{m}$  using a Yamato REM-710 microtome (Yamato Kohki Industrial Inc., Saitama, Japan) (Figure 8). The paraffin ribbons were placed in 40-50 warmed water bath and mounted onto slides. The sections were later air dried and baked  $45\text{-}50\text{ }^{\circ}\text{C}$  in an oven overnight.



**Figure 8.** Automated slide stainer. Paraffin blocks were trimmed at  $2\mu\text{m}$  using a Yamato REM-710 microtome.

### 3.3. Hematoxylin and Eosin (H&E) Staining

H&E staining is widely used in histological and cytological applications, both in fixed paraffin-embedded and frozen tissue sections. Hematoxylin produces an intense blue staining of the nuclei. Eosin stains the cytoplasm, collagen, muscle, and erythrocytes in light pink/rose. A standard H&E staining protocol includes deparaffinizing with xylene, rehydrating in decreasing concentrations of alcohol (100%, 90%, 80%, 70%, and distilled water), staining with hematoxylin, differentiating, bluing, eosin, dehydration, clearing, and coverslipping. Staining was performed using an automated slide stainer, a Tissue-Tek Prisma® Plus (Sakura Finetek Japan Co., Ltd.,) (Figure 9).



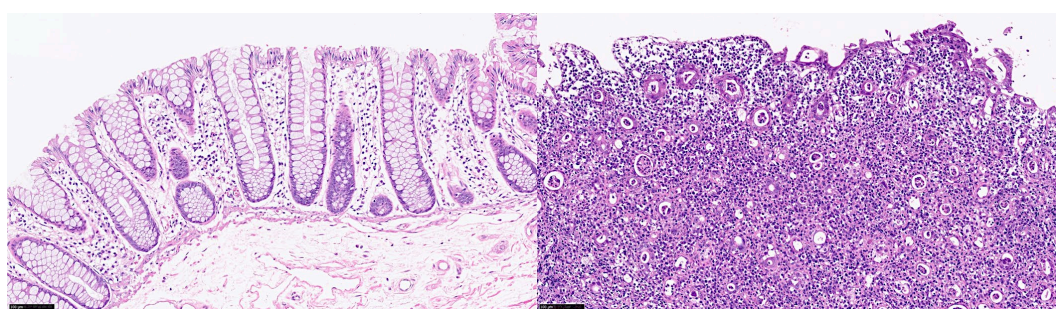
**Figure 9.** Automated slide stainer. Staining was performed using a Tissue-Tek Prisma® Plus automated slide stainer.

Glass coverslipping of histopathological slide sections was performed using a Tissue-Tek® Glas™ g2 Glass Coverslipper (Sakura Finetek Japan Co., Ltd.).

### 3.4. Score Evaluation

The endoscopic scores for chronic inflammatory bowel disease were evaluated using the Baron Score, which classifies mucosal changes into 4 grades: grade 0 (normal); grade 1 (abnormal non-hemorrhagic), 2 (moderately hemorrhagic), and 3 (severely hemorrhagic).

Ulcerative colitis histological assessment was performed using the Geboes score: grade 0 (no abnormality), grade 1 (chronic inflammatory infiltrate), grade 2 (lamina propria neutrophils and eosinophils), grade 3 (neutrophils in epithelium), and grade 4 (crypt destruction) (Figure 10).



**Figure 10.** Histologic Geboes score. Ulcerative colitis was assessed using the Geboes score, which ranges from grade 0 (no abnormality, left) to grade 4 (crypt destruction, right).

Tables A1 and A2 in the Appendix present the Baron and Geboes scores, respectively.

### 3.5. Immunohistochemistry

The immunohistochemistry of LAIR1 and TOX2 was performed using a Bond-Max fully automated immunohistochemistry and in situ hybridization staining system following the manufacturer's instructions (Leica Biosystems K.K.) (Figure 11).



**Figure 11.** Automated immunohistochemistry staining system. The immunohistochemical staining were performed using a fully automated immunohistochemistry and in situ hybridization staining system following the manufacturer's instructions (Leica BOND-MAX).

Visualization of the primary antibody bound to the tissue sections was performed using BOND Polymer Refine Detection (DS9800, Leica Biosystems K.K.). The BOND staining mode was single, i.e. a single marker and chromogen were applied to a single slide. Bond Polymer Refine Detection is a biotin-free, polymeric horseradish peroxidase (HRP)-linker antibody conjugate system for the detection of tissue bonding mouse and rabbit antibodies. This detection avoids the use of streptavidin and biotin; as a result, nonspecific staining is reduced.

The protocol sequence was the following: preparation (removal of wax using BOND Dewax Solution, 100% alcohol, and wash solution), heat-induced epitope retrieval (HIER; using BOND Epitope Retrieval ER2 solution), protein block (optional animal-free blocking solution), probe application, probe removal (wash solution), post-primary mouse linker, secondary detection (refine detection kit polymer), visualization (refine detection mixed DAB), and counterstain (hematoxylin). In summary, the staining protocol included the following steps: marker, post primary, polymer, mixed DAB, and BOND-PRIME hematoxylin.

Coverslipping was achieved using a Leica CV5030 fully automated glass coverslipper (Leica Biosystems K.K.) (Figure 12).



**Figure 12.** Glass coverslipper for immunohistochemical tissue slides. Coverslipping was achieved using a Leica CV5030 fully automated glass coverslipper.

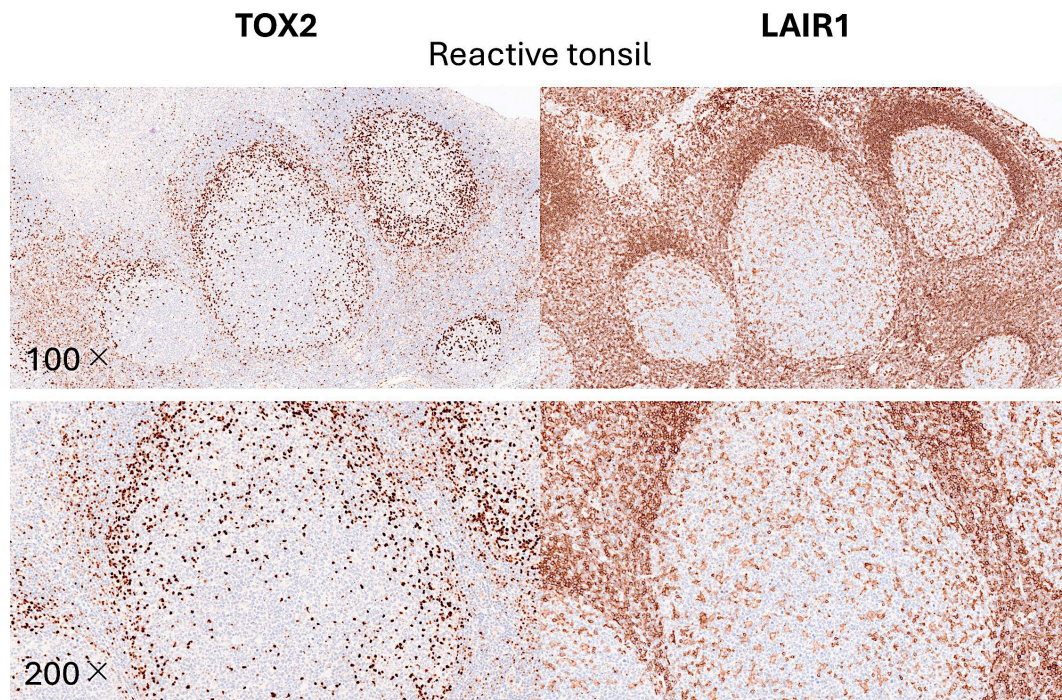
The primary antibody targeted the leukocyte-associated immunoglobulin-like receptor (LAIR1/CD305) created by the Monoclonal Antibodies Core Unit, located at the Spanish National Cancer Research Center (CNIO: Centro Nacional de Investigaciones Oncológicas; C/ Melchor Fernandez Almagro, 3, E-28029 Madrid, Spain). LAIR1 is a rat monoclonal antibody, clone JAVI82A, antigen used: RBL-1-LAIR1-MYCDDK transfected cells and last booster with LAIR1 recombinant (Gln22-His163, with a C-terminal 6-His tag); isotype IgG2a; reactivity, human; localization, membrane.

The primary antibody TOX2 targeted the TOX High Mobility Group Box Family Member 2 and was also developed by CNIO. Properties: clone name TOM924D, rat monoclonal, IgG2b K, antigen HIS-SUMO-hTOX2-Strep-tag2 full-length protein, human reactivity, nuclear localization.

Rabbit Anti-Rat IgG Antibody (H+L), Mouse Adsorbed, Unconjugated (#AI-4001-5, Vector Laboratories, Inc., Newark, CA 94560, United States) was used as a linker between primary antibodies and BOND Polymer Refine Detection.

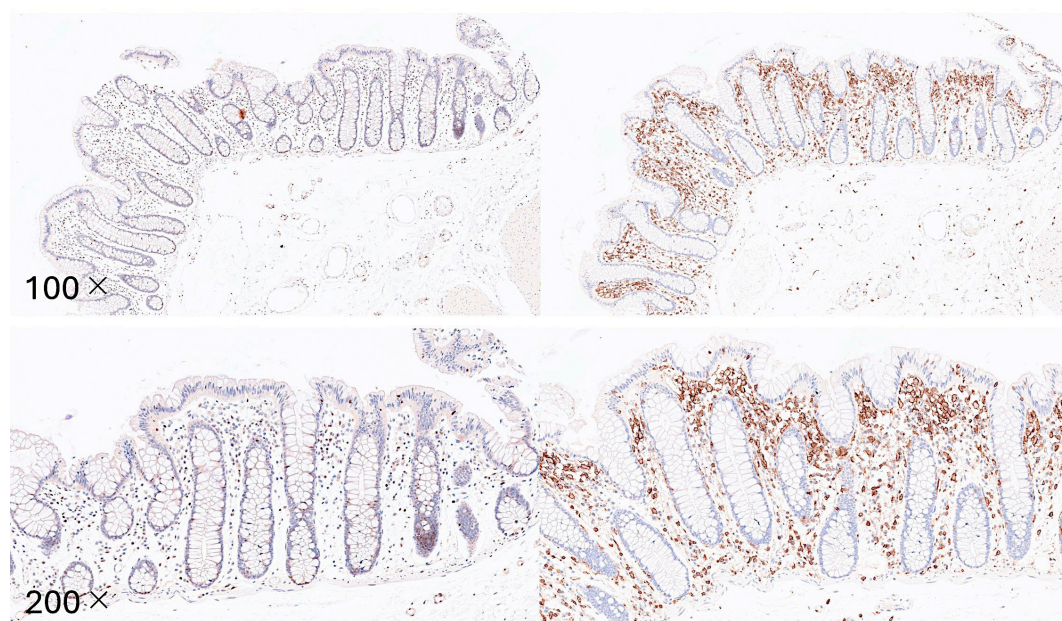
The immunohistochemistry of both antibodies was first tested in reactive lymphoid tissue (tonsils) and small and large intestinal tissue controls. Four examples of biopsies of the colon of ulcerative colitis stained with TOX2 and LAIR1 and with high infiltration of positive cells are shown in Figures 13-16.





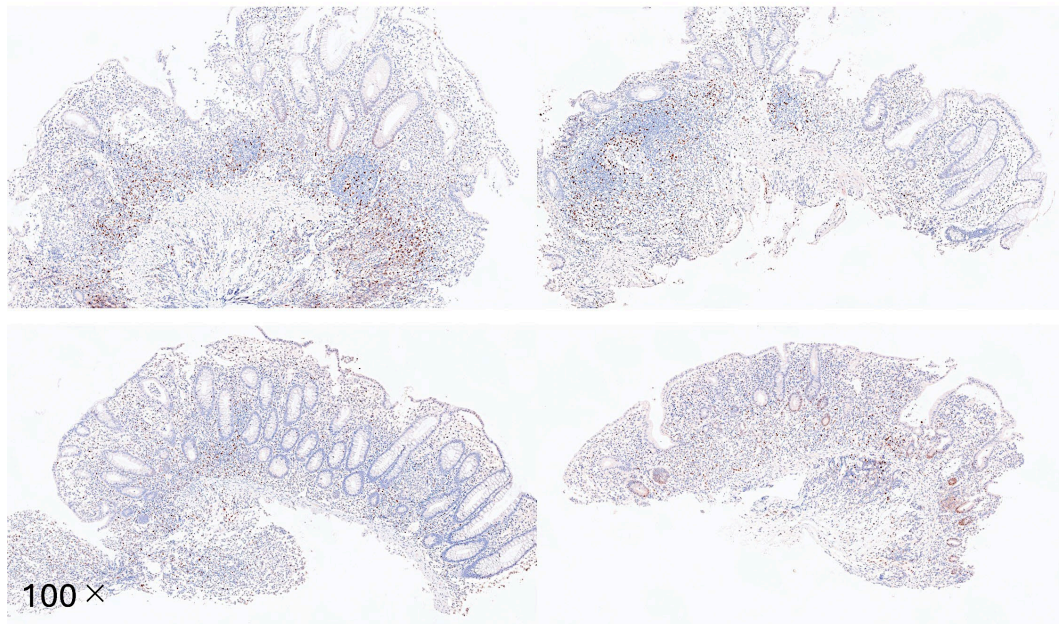
**Figure 13.** Immunohistochemistry of TOX2 and LAIR1 in reactive tonsil control. TOX2 (left column) is compatible with the staining pattern of PD-1 and in the germinal centers identified follicular T helper cells. LAIR1 is a co-inhibitory receptor found on peripheral mononuclear cells, including natural killer cells, T, and B lymphocytes. LAIR1 had a pattern compatible with macrophages/dendritic cells in the germinal centers and the interfollicular areas. Additionally, the mantle zones that include naïve B lymphocytes, were positive for LAIR1.

### Colonic mucosa control

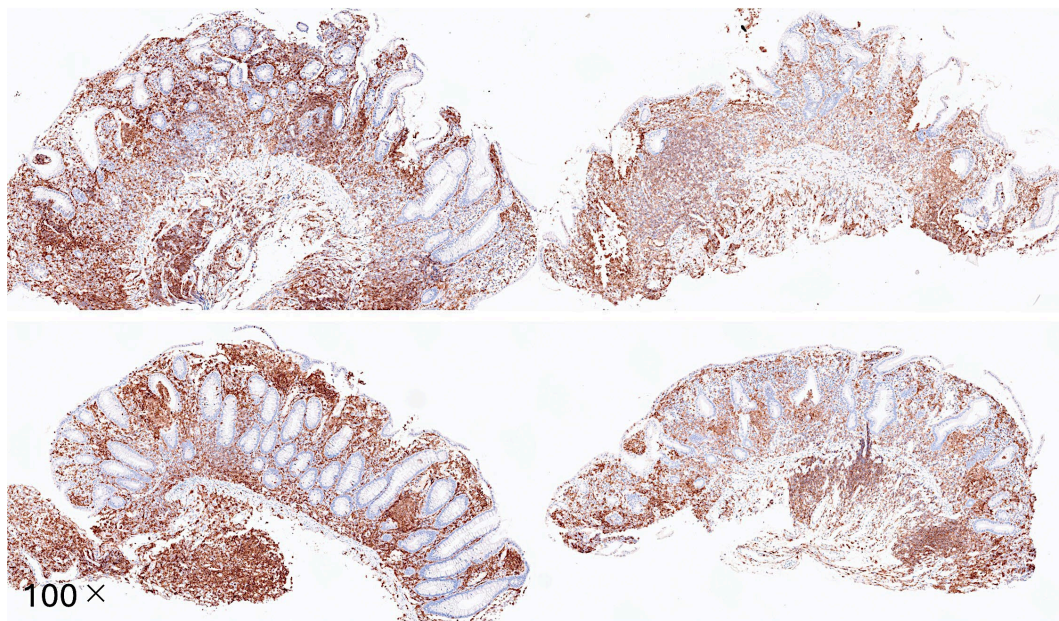


**Figure 14.** Immunohistochemistry of TOX2 and LAIR1 in colonic mucosa control. TOX2 (left column) is compatible with the staining pattern of PD-1. TOX2-positive cells were also identified in the lamina propria of the mucosa. LAIR1 (right column) is a co-inhibitory receptor found on peripheral mononuclear cells, including natural killer cells, T lymphocytes, and B lymphocytes. LAIR1 had a pattern compatible with macrophages/dendritic cells.



**Ulcerative colitis****TOX2**

**Figure 15.** Immunohistochemistry of TOX2 in ulcerative colitis. This figure shows the TOX2 staining in the mucosa of 4 biopsies. The cases had high infiltration of TOX2-positive cells in the lamina propria.

**LAIR1**

**Figure 16.** Immunohistochemistry of LAIR1 in ulcerative colitis. This figure shows LAIR1 staining in the mucosa of 4 biopsies. The patients had a high infiltration of LAIR1-positive cells in the lamina propria.

Before digitalization, the H&E and immunohistochemical slides were visualized using an Olympus upright microscope for quick evaluation and quality control purposes (Olympus BX63, Olympus K.K., Hachioji, Tokyo, Japan) (Figure 17).



**Figure 17.** Olympus upright microscope. The hematoxylin and eosin (H&E) and immunohistochemical slides were evaluated using a conventional upright microscope (Olympus BX63).

### 3.6. Whole-Slide Imaging

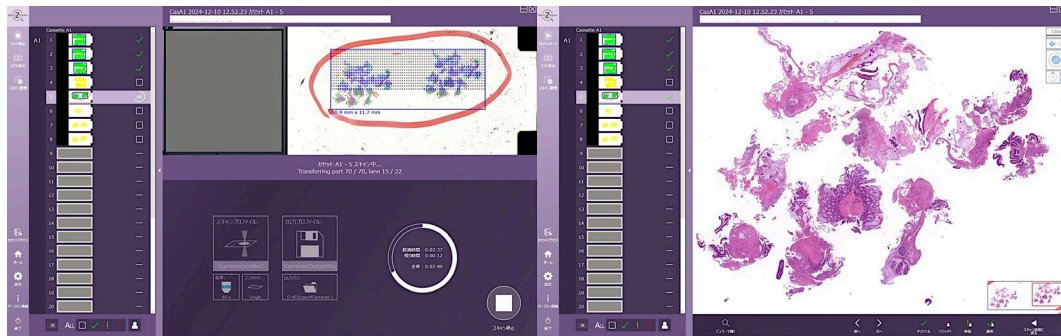
Whole-slide imaging was performed using a Hamamatsu NanoZoomer S360 scanner (Hamamatsu Photonics K.K., Hamamatsu City, 431-3196, Japan) that rapidly scanned glass slides to convert them to digital data, NZAcquire 1.2.0 software (Hamamatsu Photonics K.K.), and a Dell Precision 5820 Tower equipped with an Intel(R) Xeon(R) W-2135 CPU @ 3.70 GHz 3.70 GHz, 32.0 GB RAM, 64 bits workstation system. The operating system was Windows 10 Pro for Workstations (version 1803, build 17134.1) (Figure 18).



**Figure 18.** Whole-slide imaging. The Hamamatsu NanoZoomer S360 scanner was used to digitalize the hematoxylin and eosin (H&E) tissue samples.

The acquisition method was performed following the manufacturer's instructions. The scan was batch scan-type, at 400× magnification, and single z stack. The number of focus points was defined automatically using NZAcquire software (Figure 19).

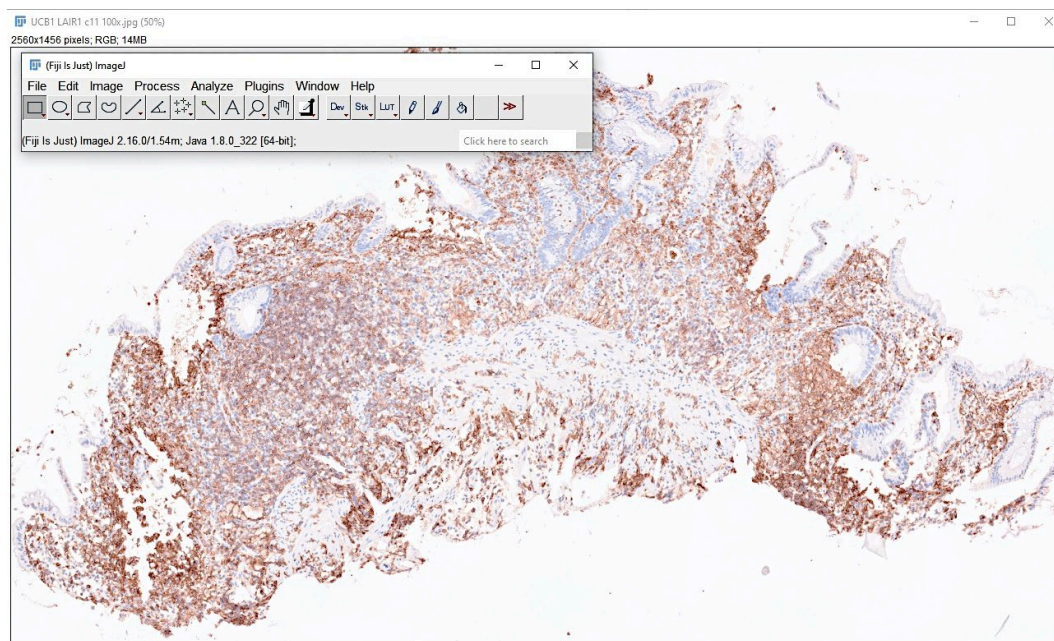




**Figure 19.** Whole-slide imaging. The glass slides were converted to digital data using a Hamamatsu NanoZoomer S360 scanner that scanned the slides at 400× magnification. .

### 3.7. Digital Image Quantification

Conventional immunohistochemical analysis was performed using digital image quantification with Fiji software, as previously described [76–79]. In summary, quantification was performed in the blue stack, min and max thresholds were set, pixels were measured, and percentages were calculated (Figure 20).



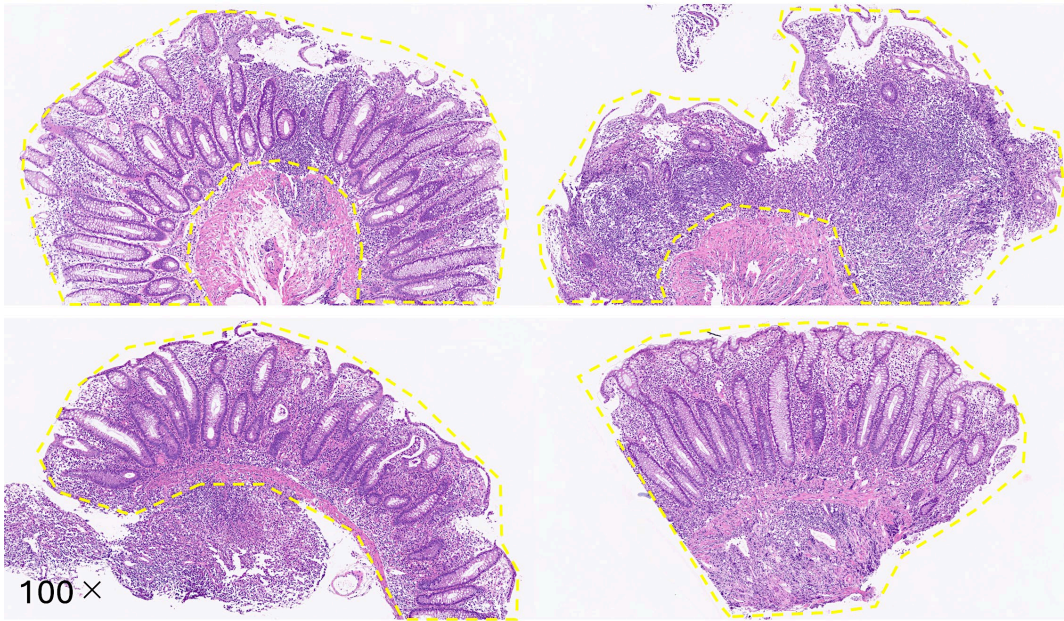
**Figure 20.** Digital image quantification. Digital image quantification was performed using Fiji software. The Blue stack was used to identify the positive staining of DAB (brown) and negative pixels.

### 3.8. Image Classification using CNNs

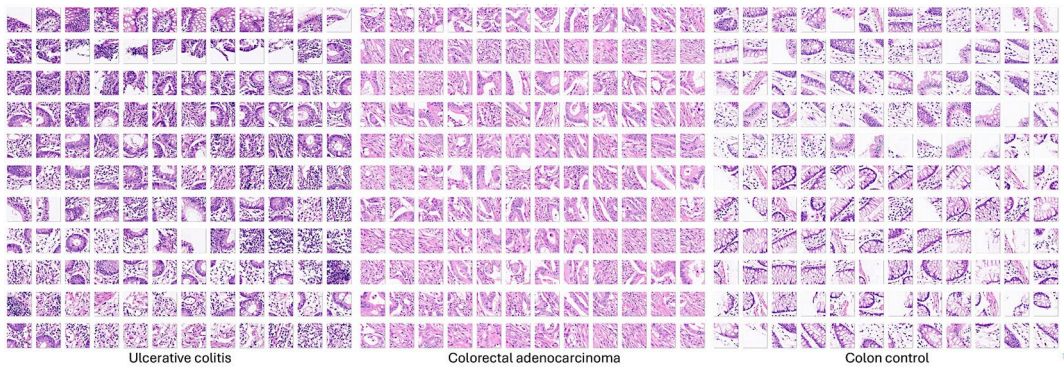
A convolutional neural network (CNN) was designed based on transfer learning from ResNet-18; and trained to classify the 3 types of image patches; ulcerative colitis (n= 9,281), colon control (n=12,246), and colorectal cancer (n=63,725). The CNN was designed in MATLAB (R2023b, update 9, 23.2.0.2668659). The CNN was also trained to differentiate between mesalazine-responsive and steroid-requiring ulcerative colitis based on H&E, LAIR1, and TOX2 staining (Figures 21 and 22).

Ulcerative colitis

H&E



**Figure 21.** Identification of regions of interest. This figure shows the hematoxylin and eosin (H&E) staining of 4 cases of endoscopic biopsy of ulcerative colitis. The areas of interest for AI analysis are indicated in yellow.



**Figure 22.** Image patches of ulcerative colitis, colorectal cancer (adenocarcinoma), and colon control.

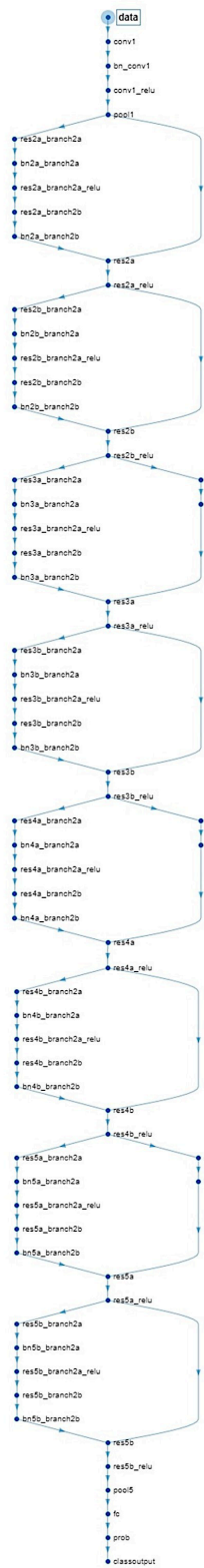
The data were partitioned into a training set (70% of the image patches) to train the network, a validation set (10%) to test the performance of the network during training, and a test set (20%) as a holdout (new data) to test the performance on new data.

The arrangement of the images was randomized to ensure that the CNN learned the classes at an even rate. Transfer learning, which involves reusing and adjusting a pretrained network, was performed on ResNet-18. For this purpose, the fully connected and classification layers of ResNet-18 were deleted and replaced with new layers. These new layers had an output size of 2. The training did not employ the augmentation technique. To avoid overfitting, the initial learning rate was set to 0.001. The maximum number of epochs was five [75].

Data normalization was applied to the input images. A detailed description of the data normalization process is provided in Appendix Table A3 [80,81].

The trained network was characterized by 71 layers. The first layer was the “ImageInputLayer”, and the last one the “ClassificationOutputLayer” (Figure 23). The parameters are listed in Table 3.





	Name	Type	Activations	Learnable Properties	States
1	data	Image Input	224(S) = 224(S) * 3(C) * 1(B)	-	-
2	conv1	2-D Convolution	112(S) = 112(S) * 64(C) * 1(B)	Weights 7 * 7 * 3 = 64 Bias 1 * 1 * 64	-
3	bn_conv1	Batch Normalization	112(S) = 112(S) * 64(C) * 1(B)	Offset 1 * 1 * 64 Scale 1 * 1 * 64	TrainedMean 1 * 1 * 64 TrainedVariance 1 * 1 * 64
4	conv1_relu	ReLU	112(S) = 112(S) * 64(C) * 1(B)	-	-
5	pool1	2-D Max Pooling	56(S) = 56(S) * 64(C) * 1(B)	-	-
6	res2a_branch2a	2-D Convolution	56(S) = 56(S) * 64(C) * 1(B)	Weights 3 * 3 * 64 = 64 Bias 1 * 1 * 64	-
7	bn2a_branch2a	Batch Normalization	56(S) = 56(S) * 64(C) * 1(B)	Offset 1 * 1 * 64 Scale 1 * 1 * 64	TrainedMean 1 * 1 * 64 TrainedVariance 1 * 1 * 64
8	res2a_branch2a_relu	ReLU	56(S) = 56(S) * 64(C) * 1(B)	-	-
9	res2a_branch2b	2-D Convolution	56(S) = 56(S) * 64(C) * 1(B)	Weights 3 * 3 * 64 = 64 Bias 1 * 1 * 64	-
10	bn2a_branch2b	Batch Normalization	56(S) = 56(S) * 64(C) * 1(B)	Offset 1 * 1 * 64 Scale 1 * 1 * 64	TrainedMean 1 * 1 * 64 TrainedVariance 1 * 1 * 64
11	res2a	Element-wise addition of 2 inputs	56(S) = 56(S) * 64(C) * 1(B)	-	-
12	res2a_relu	ReLU	56(S) = 56(S) * 64(C) * 1(B)	-	-
13	res2b_branch2a	2-D Convolution	56(S) = 56(S) * 64(C) * 1(B)	Weights 3 * 3 * 64 = 64 Bias 1 * 1 * 64	-
14	bn2b_branch2a	Batch Normalization	56(S) = 56(S) * 64(C) * 1(B)	Offset 1 * 1 * 64 Scale 1 * 1 * 64	TrainedMean 1 * 1 * 64 TrainedVariance 1 * 1 * 64
15	res2b_branch2a_relu	ReLU	56(S) = 56(S) * 64(C) * 1(B)	-	-
16	res2b_branch2b	2-D Convolution	56(S) = 56(S) * 64(C) * 1(B)	Weights 3 * 3 * 64 = 64 Bias 1 * 1 * 64	-
17	bn2b_branch2b	Batch Normalization	56(S) = 56(S) * 64(C) * 1(B)	Offset 1 * 1 * 64 Scale 1 * 1 * 64	TrainedMean 1 * 1 * 64 TrainedVariance 1 * 1 * 64
18	res2b	Element-wise addition of 2 inputs	56(S) = 56(S) * 64(C) * 1(B)	-	-
19	res2b_relu	ReLU	56(S) = 56(S) * 64(C) * 1(B)	-	-
20	res2a_branch2a	2-D Convolution	28(S) = 28(S) * 128(C) * 1(B)	Weights 3 * 3 * 64 = 128 Bias 1 * 1 * 128	-
21	bn2a_branch2a	Batch Normalization	28(S) = 28(S) * 128(C) * 1(B)	Offset 1 * 1 * 128 Scale 1 * 1 * 128	TrainedMean 1 * 1 * 128 TrainedVariance 1 * 1 * 128
22	res2a_branch2a_relu	ReLU	28(S) = 28(S) * 128(C) * 1(B)	-	-
23	res2a_branch2b	2-D Convolution	28(S) = 28(S) * 128(C) * 1(B)	Weights 3 * 3 * 128 = 128 Bias 1 * 1 * 128	-
24	bn2a_branch2b	Batch Normalization	28(S) = 28(S) * 128(C) * 1(B)	Offset 1 * 1 * 128 Scale 1 * 1 * 128	TrainedMean 1 * 1 * 128 TrainedVariance 1 * 1 * 128
25	res2a_branch2b_relu	ReLU	28(S) = 28(S) * 128(C) * 1(B)	-	-
26	res2a_branch1	2-D Convolution	28(S) = 28(S) * 128(C) * 1(B)	Weights 1 * 1 * 64 = 128 Bias 1 * 1 * 128	-
27	bn2a_branch1	Batch Normalization	28(S) = 28(S) * 128(C) * 1(B)	Offset 1 * 1 * 128 Scale 1 * 1 * 128	TrainedMean 1 * 1 * 128 TrainedVariance 1 * 1 * 128
28	res2a	Element-wise addition of 2 inputs	28(S) = 28(S) * 128(C) * 1(B)	-	-
29	res2a_relu	ReLU	28(S) = 28(S) * 128(C) * 1(B)	-	-
30	res2b_branch2a	2-D Convolution	28(S) = 28(S) * 128(C) * 1(B)	Weights 3 * 3 * 128 = 128 Bias 1 * 1 * 128	-
31	bn2b_branch2a	Batch Normalization	28(S) = 28(S) * 128(C) * 1(B)	Offset 1 * 1 * 128 Scale 1 * 1 * 128	TrainedMean 1 * 1 * 128 TrainedVariance 1 * 1 * 128
32	res2b_branch2a_relu	ReLU	28(S) = 28(S) * 128(C) * 1(B)	-	-
33	res2b_branch2b	2-D Convolution	28(S) = 28(S) * 128(C) * 1(B)	Weights 3 * 3 * 128 = 128 Bias 1 * 1 * 128	-
34	bn2b_branch2b	Batch Normalization	28(S) = 28(S) * 128(C) * 1(B)	Offset 1 * 1 * 128 Scale 1 * 1 * 128	TrainedMean 1 * 1 * 128 TrainedVariance 1 * 1 * 128
35	res2b	Element-wise addition of 2 inputs	28(S) = 28(S) * 128(C) * 1(B)	-	-
36	res2b_relu	ReLU	28(S) = 28(S) * 128(C) * 1(B)	-	-
37	res2a_branch2a	2-D Convolution	14(S) = 14(S) * 256(C) * 1(B)	Weights 3 * 3 * 128 = 256 Bias 1 * 1 * 256	-
38	bn2a_branch2a	Batch Normalization	14(S) = 14(S) * 256(C) * 1(B)	Offset 1 * 1 * 256 Scale 1 * 1 * 256	TrainedMean 1 * 1 * 256 TrainedVariance 1 * 1 * 256
39	res2a_branch2a_relu	ReLU	14(S) = 14(S) * 256(C) * 1(B)	-	-
40	res2a_branch2b	2-D Convolution	14(S) = 14(S) * 256(C) * 1(B)	Weights 3 * 3 * 256 = 256 Bias 1 * 1 * 256	-
41	bn2a_branch2b	Batch Normalization	14(S) = 14(S) * 256(C) * 1(B)	Offset 1 * 1 * 256 Scale 1 * 1 * 256	TrainedMean 1 * 1 * 256 TrainedVariance 1 * 1 * 256
42	res2a_branch2b_relu	ReLU	14(S) = 14(S) * 256(C) * 1(B)	-	-
43	res2a_branch1	2-D Convolution	14(S) = 14(S) * 256(C) * 1(B)	Weights 1 * 1 * 128 = 256 Bias 1 * 1 * 256	-
44	bn2a_branch1	Batch Normalization	14(S) = 14(S) * 256(C) * 1(B)	Offset 1 * 1 * 256 Scale 1 * 1 * 256	TrainedMean 1 * 1 * 256 TrainedVariance 1 * 1 * 256
45	res2a	Element-wise addition of 2 inputs	14(S) = 14(S) * 256(C) * 1(B)	-	-
46	res2a_relu	ReLU	14(S) = 14(S) * 256(C) * 1(B)	-	-
47	res2b_branch2a	2-D Convolution	14(S) = 14(S) * 256(C) * 1(B)	Weights 3 * 3 * 256 = 256 Bias 1 * 1 * 256	-
48	bn2b_branch2a	Batch Normalization	14(S) = 14(S) * 256(C) * 1(B)	Offset 1 * 1 * 256 Scale 1 * 1 * 256	TrainedMean 1 * 1 * 256 TrainedVariance 1 * 1 * 256
49	res2b_branch2a_relu	ReLU	14(S) = 14(S) * 256(C) * 1(B)	-	-
50	res2b_branch2b	2-D Convolution	14(S) = 14(S) * 256(C) * 1(B)	Weights 3 * 3 * 256 = 256 Bias 1 * 1 * 256	-
51	bn2b_branch2b	Batch Normalization	14(S) = 14(S) * 256(C) * 1(B)	Offset 1 * 1 * 256 Scale 1 * 1 * 256	TrainedMean 1 * 1 * 256 TrainedVariance 1 * 1 * 256
52	res2b	Element-wise addition of 2 inputs	14(S) = 14(S) * 256(C) * 1(B)	-	-
53	res2b_relu	ReLU	14(S) = 14(S) * 256(C) * 1(B)	-	-
54	res2a_branch2a	2-D Convolution	7(S) = 7(S) * 512(C) * 1(B)	Weights 3 * 3 * 256 = 512 Bias 1 * 1 * 512	-
55	bn2a_branch2a	Batch Normalization	7(S) = 7(S) * 512(C) * 1(B)	Offset 1 * 1 * 512 Scale 1 * 1 * 512	TrainedMean 1 * 1 * 512 TrainedVariance 1 * 1 * 512
56	res2a_branch2a_relu	ReLU	7(S) = 7(S) * 512(C) * 1(B)	-	-
57	res2a_branch2b	2-D Convolution	7(S) = 7(S) * 512(C) * 1(B)	Weights 3 * 3 * 512 = 512 Bias 1 * 1 * 512	-
58	bn2a_branch2b	Batch Normalization	7(S) = 7(S) * 512(C) * 1(B)	Offset 1 * 1 * 512 Scale 1 * 1 * 512	TrainedMean 1 * 1 * 512 TrainedVariance 1 * 1 * 512
59	res2a_branch2b_relu	ReLU	7(S) = 7(S) * 512(C) * 1(B)	-	-
60	res2a_branch1	2-D Convolution	7(S) = 7(S) * 512(C) * 1(B)	Weights 1 * 1 * 256 = 512 Bias 1 * 1 * 512	-
61	bn2a_branch1	Batch Normalization	7(S) = 7(S) * 512(C) * 1(B)	Offset 1 * 1 * 512 Scale 1 * 1 * 512	TrainedMean 1 * 1 * 512 TrainedVariance 1 * 1 * 512
62	res2a	Element-wise addition of 2 inputs	7(S) = 7(S) * 512(C) * 1(B)	-	-
63	res2a_relu	ReLU	7(S) = 7(S) * 512(C) * 1(B)	-	-
64	res2b_branch2a	2-D Convolution	7(S) = 7(S) * 512(C) * 1(B)	Weights 3 * 3 * 512 = 512 Bias 1 * 1 * 512	-
65	bn2b_branch2a	Batch Normalization	7(S) = 7(S) * 512(C) * 1(B)	Offset 1 * 1 * 512 Scale 1 * 1 * 512	TrainedMean 1 * 1 * 512 TrainedVariance 1 * 1 * 512
66	res2b_branch2a_relu	ReLU	7(S) = 7(S) * 512(C) * 1(B)	-	-
67	res2b_branch2b	2-D Convolution	7(S) = 7(S) * 512(C) * 1(B)	Weights 3 * 3 * 512 = 512 Bias 1 * 1 * 512	-
68	bn2b_branch2b	Batch Normalization	7(S) = 7(S) * 512(C) * 1(B)	Offset 1 * 1 * 512 Scale 1 * 1 * 512	TrainedMean 1 * 1 * 512 TrainedVariance 1 * 1 * 512
69	res2b	Element-wise addition of 2 inputs	7(S) = 7(S) * 512(C) * 1(B)	-	-
70	res2b_relu	ReLU	7(S) = 7(S) * 512(C) * 1(B)	-	-
71	pool5	2-D Global Average...	1(S) = 1(S) * 512(C) * 1(B)	-	-
72	fc	Fully Connected	1(S) = 1(S) * 3(C) * 1(B)	Weights 3 * 512 Bias 3 * 1	-
73	prob	Softmax	1(S) = 1(S) * 3(C) * 1(B)	-	-
74	classoutput	Classification Output	1(S) = 1(S) * 3(C) * 1(B)	-	-

Figure 23. Trained network layers. The CNN was designed using a transfer learning strategy and ResNet-18.

The code used is shown in Table A4 in the Appendix

Table 3. Design and training parameters.

ResNet-18-Based CNN	Training (70%)	Validation (10%)	Training Options
---------------------	----------------	------------------	------------------



Input type: image patches	Observations: 59,677	Observations: 8525	Solver: sgdm
Output type: classification	Classes: 3	Classes: 3	Initial learning rate: 0.001
Number of layers: 71	Ulcerative colitis: 6497	Ulcerative colitis: 928	MiniBatch size: 128
Number of connections: 78	Colorectal cancer: 44,608	Colorectal cancer: 6372	MaxEpochs: 5
	Colon control: 8572	Colon control: 1225	Validation frequency: 50
			Iterations: 2330
			Iterations per epoch: 466

Table 4. Additional training parameters.

<b>Additional detailed training options</b>
<b>Import images</b>
Augmentation options: none
Available parameters
Random reflection axis: x, y
Random rotation (degrees): min, max
Random rescaling: min, max
Random horizontal translation (pixels): min, max
Random vertical translation (pixels): min, max
Resize during training to match network input size: yes, no
<b>Solver</b>
Momentum: 0.9
<b>Learn rate</b>
LearnRateSchedule: none
LearnRateDropFactor: 0.1
LearnRateDropPeriod: 10
<b>Normalization and Regularization</b>
L2Regularization: 0.0001
ResetInputNormalization: yes
BatchNormalizationStatistics: population
<b>Mini-Batch</b>
Shuffle: every-epoch
<b>Validation and Output</b>
ValidationPatience: Inf
OutputNetwork: last-iteration
<b>Gradient Clipping</b>
GradientThresholdMethod: l2norm
GradientThreshold: Inf
<b>Hardware</b>
ExecutionThreshold: auto.
<b>Checkpoint</b>
CheckpointPath: n/a
CheckpointFrequency: 1
CheckpointFrequencyUnit: epoch

Based on ResNet-18 transfer learning. Convolutional neural network, CNN.

In the analysis setup, the image patches of the 3 diagnoses were pooled in 3 different folders. The content of each folder was then split into training, validation, and testing sets. As a result, no image patches were repeated in different folders. Nonetheless, some researchers believe that this strategy could lead to information leakage. Therefore, an additional and independent test set of 10 cases of colorectal cancer (adenocarcinoma) was used to confirm the performance of the trained CNN. Each patient was analyzed independently [75].

Of note, other types of CNNs were tested in this study. The LAIR1 and TOX2 biomarkers were included in the CNN training in an independent analysis of H&E staining. A differentiation between

steroid-requiring (SR) and mesalazine-responsive ulcerative colitis using LAIR1 and TOX2 immunohistochemistry.

The performance parameters were the following: accuracy, precision, recall, F1 Score, specificity, and false positive rate.

The performance of ResNet-18 was compared to that of other CNNs under the same experimental conditions, including DenseNet-201, ResNet-50, Inception-v3, ResNet-101, ShuffleNet, MobileNet-v2, NasNet-Large, GoogLeNet-Places365, VGG-19, EfficientNet-b0, AlexNet, Xception, VGG-16, GoogLeNet, and NasNet-Mobile.

### 3.9. Computational Requirements

All analyses were performed using a desktop computer equipped with an AMD Ryzen 9 7950X CPU, 128 Gb of RAM (Crucial Desktop DDR5-4800 UDIMM 1.1V CL40, CT2K32G48C40U5 x2), and an Nvidia GeForce RTX 4090 graphics card (ASUS ROG Strix GeForce RTX® 4090 OC Edition 24GB GDDR6X). The specifications of the GPU were 16384 Nvidia Cuda® cores, 4<sup>th</sup> Generation 1321 AI TOPS Tensor cores, 2.23 GHz base and 2.52 GHz boost clock, 24 GB GDDR6X memory, 384-bit memory interface width, Ada Lovelace architecture (Figure 24).



**Figure 24.** Computational requirements. The desktop workstation was equipped with a Nvidia GeForce RTX 4090 graphics card.

## 4. Results

This paper of Methods type is a companion manuscript to a recently published article entitled “Ulcerative Colitis, LAIR1 and TOX2 Expression, and Colorectal Cancer Deep Learning Image Classification Using Convolutional Neural Networks”. Published in *Cancers (Basel)*; Published 2024 Dec 19. doi:10.3390/cancers16244230 [75].

In summary, steroid-requiring ulcerative colitis was characterized by higher endoscopic Baron and histologic Geboes scores, higher LAIR1 immunohistochemical expression in the lamina propria, and lower TOX2 in the isolated lymphoid follicles (ILFs). The CNN successfully classified ulcerative colitis, colorectal cancer, and colon control with high performance. The classification was also validated using an independent test sample.

## 5. Discussion

Artificial Intelligence (AI) is the science and engineering that allows computers and machines to simulate human learning and problem solving. The development of AI follows a series of steps since 1950s, when the term AI was first postulated as human intelligence exhibited by machines. In the 1980s, machine learning was developed as an AI system that could learn from historical data. In the 2010s, deep learning was introduced as a machine learning model that simulated human brain function. From 2020 generative AI (Gen AI) was developed as a deep learning model (foundation models) that creates original content [82–84].

There are many techniques or algorithms of machine learning such as logistic regression, decision trees, random forest, and k-nearest neighbor (KNN) [79,80]. Deep learning is based on neural networks (deep neural networks) that are designed as input, hidden, and output layers. Deep learning enables semi-supervised, self-supervised, reinforcement, and transfer learning [82–84].

This technical note manuscript described a dataset and computer vision method that allows software and machines to understand and classify visual data in the context of artificial narrow AI. Narrow AI is also referred to as Weak AI, and it is the only type of AI that exists today. Narrow AI can be trained to perform a single task, and it can perform faster and better than the human mind. However, it cannot perform tasks outside its defined scope.

6. Conclusions

Convolutional neural networks are suitable for image classification in the context of narrow artificial intelligence. This paper describes a technical note on a method for collecting, processing, managing, and analyzing scientific data from intestinal histological images. Related source code is provided in the Appendix, and examples of image patches are provided in the Supplementary Material.

**Supplementary Materials:** Image\_patches\_examples\_datav2-3439728.

**Author Contributions:** Conceptualization, J.C.; methodology, J.C., G.R. and R.H.; software, J.C.; formal analysis, J.C.; primary antibodies, J.C. and G.R.; writing—original draft preparation, J.C.; writing—review, J.C. and R.H.; funding acquisition, J.C. and R.H. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Ministry of Education, Culture, Sports, Science, and Technology of Japan; and KAKEN grants 23K06454, 18K15100, and 15K19061. R.H. is funded by the University of Sharjah (grant no: 24010902153) and ASPIRE, the technology program management pillar of Abu Dhabi’s Advanced Technology Research Council (ATRC), via the ASPIRE Precision Medicine Research Institute Abu Dhabi (VRI-20–10).

**Institutional Review Board Statement:** The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (Ethics Committee) of TOKAI UNIVERSITY, SCHOOL OF MEDICINE (protocol code IRB14R-080, IRB20-156, and 13R-119).

**Informed Consent Statement:** Informed consent was obtained from all participants.

**Data Availability Statement:** All data are available from Zenodo CERN and OpenAIRE Open Science repository: Carreras, J. (2024). Image dataset (Version 1) [Data set]. Zenodo. <https://doi.org/10.5281/zenodo.14429385>. All raw data and methodology are available upon request from Joaquim Carreras ([joaquim.carreras@tokai.ac.jp](mailto:joaquim.carreras@tokai.ac.jp)).

**Acknowledgments:** None.

**Conflicts of Interest:** The authors declare no conflict of interest.

Appendix Table A

Table A1. Endoscopic Baron score.

0	Normal: matte mucosa, ramifying vascular pattern clearly visible, no spontaneous bleeding, no bleeding to light touch.
1	Abnormal, but non-hemorrhagic: appearance between 0 and 2.
2	Moderately hemorrhagic: bleeding to light touch, but no spontaneous bleeding ahead of the instrument on initial inspection
3	Severely hemorrhagic: spontaneous bleeding ahead of instrument at initial inspection and bleeding to light touch

Table A2. Histologic Geboes score.

Grade 0	Structural (architectural changes)
---------	------------------------------------



Subgrades	
0	No abnormality
0.1	Mild abnormality
0.2	Mild or moderate diffuse or multifocal abnormalities
0.3	Severe diffuse or multifocal abnormalities
<b>Grade 1</b>	<b>Chronic inflammatory infiltrate</b>
Subgrades	
1	No increase
1.1	Mild but unequivocal increase
1.2	Moderate increase
1.3	Marked increase
<b>Grade 2</b>	<b>Lamina propria neutrophils and eosinophils</b>
<i>2A Eosinophils</i>	
2A.0	No increase
2A.1	Mild but unequivocal increase
2A.2	Moderate increase
2A.3	Marked increase
<i>2B Neutrophils</i>	
2B.0	No increase
2B.1	Mild but unequivocal increase
2B.2	Moderate increase
2B.3	Marked increase
<b>Grade 3</b>	<b>Neutrophils in epithelium</b>
<i>Subgrades</i>	
3.0	None
3.1	<5% Crypts involves
3.2	<50% Crypts involves
3.3	>50% Crypts involves
<b>Grade 4</b>	<b>Crypt destruction</b>
<i>Subgrades</i>	
4.0	None
4.1	Probable — local excess of neutrophils in part of crypt
4.2	Probable — marked attenuation
4.3	Unequivocal crypt destruction
<b>Grade 5</b>	<b>Erosion or ulceration</b>
<i>Subgrades</i>	
5.0	No erosion, ulceration, or granulation tissue
5.1	Recovering epithelium + adjacent inflammation
5.2	Probable erosion focally stripped
5.3	Unequivocal erosion
5.4	Ulcer or granulation tissue

**Table A3.** Data normalization.

Data normalization was applied to the input images: `imageInputLayer` (an image input layer inputs 2-D images to a neural network and applies data normalization), and `batchNormalizationLayer` (a batch normalization layer independently normalizes a mini-batch of data across all observations for each channel. To accelerate the training of the CNN and reduce the sensitivity to network initialization, batch normalization layers are used between the convolutional layers and nonlinearities, such as ReLU layers. `Layer = batchNormalizationLayer (Name, Value)` creates a batch normalization layer and sets the optional `TrainedMean`, `TrainedVariance`, `Epsilon`, `Parameters` and `Initialization`, `Learning Rate` and `Regularization`, and `Name` properties using one or more name-value pairs. After normalization, the layer scales the input with a learnable scale factor  $\gamma$  and shifts it by a learnable offset  $\beta$ ) [80,81]

**Table A4.** Code.

**Ⓐ To load training setup data:**

```
trainingSetup = load("...")
```

**Ⓑ Import data:**

```
imdsTrain = trainingSetup.imdsTrain;
imdsValidation = trainingSetup.imdsValidation;
```

**Ⓒ Resize the images to match the network input layer:**

```
augimdsTrain = augmentedImageDatastore([224 224 3],imdsTrain);
augimdsValidation = augmentedImageDatastore([224 224 3],imdsValidation);
```

**Ⓓ Set training options:**

```
opts = trainingOptions("sgdm",...
    "ExecutionEnvironment","auto",...
    "InitialLearnRate",0.001,...
    "MaxEpochs",5,...
    "Shuffle","every-epoch",...
    "Plots","training-progress",...
    "ValidationData",augimdsValidation);
```

**Ⓔ Create layer graph:**

```
lgraph = layerGraph();
```

**Ⓕ Add layer branches:**

```
tempLayers = [
    imageInputLayer([224 224 3],"Name","data","Normalization","zscore","Mean",trainingSetup.data.Mean,"StandardDeviation",
    trainingSetup.data.StandardDeviation)
    convolution2dLayer([7 7],64,"Name","conv1","BiasLearnRateFactor",0,"Padding",[3 3 3],
    "Stride",[2 2],"Bias",trainingSetup.conv1.Bias,"Weights",trainingSetup.conv1.Weights)
```

```
batchNormalizationLayer("Name","bn_conv1","Offset",trainingSetup.bn_conv1.Offset,"Scale",trainingSetup.bn_conv1.Scale,"TrainedMean",trainingSetup.bn_conv1.TrainedMean,"TrainedVariance",trainingSetup.bn_conv1.TrainedVariance)
```

```
reluLayer("Name","conv1_relu")
maxPooling2dLayer([3 3],"Name","pool1","Padding",[1 1 1],"Stride",[2 2]);
lgraph = addLayers(lgraph,tempLayers);
```

```
tempLayers = [
    convolution2dLayer([3 3],64,"Name","res2a_branch2a","BiasLearnRateFactor",0,"Padding",[1 1 1],"Bias",trainingSetup.res2a_branch2a.Bias,"Weights",trainingSetup.res2a_branch2a.Weights)
```

```
batchNormalizationLayer("Name","bn2a_branch2a","Offset",trainingSetup.bn2a_branch2a.Offset,"Scale",trainingSetup.bn2a_branch2a.Scale,"TrainedMean",trainingSetup.bn2a_branch2a.TrainedMean,"TrainedVariance",trainingSetup.bn2a_branch2a.TrainedVariance)
```

```
reluLayer("Name","res2a_branch2a_relu")
convolution2dLayer([3 3],64,"Name","res2a_branch2b","BiasLearnRateFactor",0,"Padding",[1 1 1],"Bias",trainingSetup.res2a_branch2b.Bias,"Weights",trainingSetup.res2a_branch2b.Weights)
```

```
batchNormalizationLayer("Name","bn2a_branch2b","Offset",trainingSetup.bn2a_branch2b.Offset,"Scale",trainingSetup.bn2a_branch2b.Scale,"TrainedMean",trainingSetup.bn2a_branch2b.TrainedMean,"TrainedVariance",trainingSetup.bn2a_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);
```

```
tempLayers = [
    additionLayer(2,"Name","res2a")
    reluLayer("Name","res2a_relu");
lgraph = addLayers(lgraph,tempLayers);
```

```
tempLayers = [
    convolution2dLayer([3 3],64,"Name","res2b_branch2a","BiasLearnRateFactor",0,"Padding",[1 1 1],"Bias",trainingSetup.res2b_branch2a.Bias,"Weights",trainingSetup.res2b_branch2a.Weights)
```

```
batchNormalizationLayer("Name","bn2b_branch2a","Offset",trainingSetup.bn2b_branch2a.Offset,
```

---

```

"Scale",trainingSetup.bn2b_branch2a.Scale,"TrainedMean",trainingSetup.bn2b_branch2a.Trained
Mean,"TrainedVariance",trainingSetup.bn2b_branch2a.TrainedVariance)
    reluLayer("Name","res2b_branch2a_relu")
    convolution2dLayer([3 3],64,"Name","res2b_branch2b","BiasLearnRateFactor",0,"Padding",[1
1 1 1],"Bias",trainingSetup.res2b_branch2b.Bias,"Weights",trainingSetup.res2b_branch2b.Weights)

batchNormalizationLayer("Name","bn2b_branch2b","Offset",trainingSetup.bn2b_branch2b.Offset
,"Scale",trainingSetup.bn2b_branch2b.Scale,"TrainedMean",trainingSetup.bn2b_branch2b.Trained
Mean,"TrainedVariance",trainingSetup.bn2b_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    additionLayer(2,"Name","res2b")
    reluLayer("Name","res2b_relu")];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([3
3],128,"Name","res3a_branch2a","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res3a_branch2a.Bias,"Weights",trainingSetup.res3a_branch2a.Weights)

batchNormalizationLayer("Name","bn3a_branch2a","Offset",trainingSetup.bn3a_branch2a.Offset,
"Scale",trainingSetup.bn3a_branch2a.Scale,"TrainedMean",trainingSetup.bn3a_branch2a.Trained
Mean,"TrainedVariance",trainingSetup.bn3a_branch2a.TrainedVariance)
    reluLayer("Name","res3a_branch2a_relu")
    convolution2dLayer([3
3],128,"Name","res3a_branch2b","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res3a_branch2b.Bias,"Weights",trainingSetup.res3a_branch2b.Weights)

batchNormalizationLayer("Name","bn3a_branch2b","Offset",trainingSetup.bn3a_branch2b.Offset,
"Scale",trainingSetup.bn3a_branch2b.Scale,"TrainedMean",trainingSetup.bn3a_branch2b.Trained
Mean,"TrainedVariance",trainingSetup.bn3a_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([1 1],128,"Name","res3a_branch1","BiasLearnRateFactor",0,"Stride",[2
2],"Bias",trainingSetup.res3a_branch1.Bias,"Weights",trainingSetup.res3a_branch1.Weights)

batchNormalizationLayer("Name","bn3a_branch1","Offset",trainingSetup.bn3a_branch1.Offset,"S
cale",trainingSetup.bn3a_branch1.Scale,"TrainedMean",trainingSetup.bn3a_branch1.TrainedMea
n,"TrainedVariance",trainingSetup.bn3a_branch1.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    additionLayer(2,"Name","res3a")
    reluLayer("Name","res3a_relu")];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([3
3],128,"Name","res3b_branch2a","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res3b_branch2a.Bias,"Weights",trainingSetup.res3b_branch2a.Weights)

batchNormalizationLayer("Name","bn3b_branch2a","Offset",trainingSetup.bn3b_branch2a.Offset,
"Scale",trainingSetup.bn3b_branch2a.Scale,"TrainedMean",trainingSetup.bn3b_branch2a.Trained
Mean,"TrainedVariance",trainingSetup.bn3b_branch2a.TrainedVariance)
    reluLayer("Name","res3b_branch2a_relu")
    convolution2dLayer([3
3],128,"Name","res3b_branch2b","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res3b_branch2b.Bias,"Weights",trainingSetup.res3b_branch2b.Weights)

batchNormalizationLayer("Name","bn3b_branch2b","Offset",trainingSetup.bn3b_branch2b.Offset
,"Scale",trainingSetup.bn3b_branch2b.Scale,"TrainedMean",trainingSetup.bn3b_branch2b.Trained
Mean,"TrainedVariance",trainingSetup.bn3b_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

```

---



---

```

tempLayers = [
    additionLayer(2,"Name","res3b")
    reluLayer("Name","res3b_relu");
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([3
3],256,"Name","res4a_branch2a","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res4a_branch2a.Bias,"Weights",trainingSetup.res4a_branch2a.Weights)

    batchNormalizationLayer("Name","bn4a_branch2a","Offset",trainingSetup.bn4a_branch2a.Offset,
"Scale",trainingSetup.bn4a_branch2a.Scale,"TrainedMean",trainingSetup.bn4a_branch2a.Trained
Mean,"TrainedVariance",trainingSetup.bn4a_branch2a.TrainedVariance)
    reluLayer("Name","res4a_branch2a_relu")
    convolution2dLayer([3
3],256,"Name","res4a_branch2b","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res4a_branch2b.Bias,"Weights",trainingSetup.res4a_branch2b.Weights)

    batchNormalizationLayer("Name","bn4a_branch2b","Offset",trainingSetup.bn4a_branch2b.Offset,
"Scale",trainingSetup.bn4a_branch2b.Scale,"TrainedMean",trainingSetup.bn4a_branch2b.Trained
Mean,"TrainedVariance",trainingSetup.bn4a_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([1 1],256,"Name","res4a_branch1","BiasLearnRateFactor",0,"Stride",[2
2],"Bias",trainingSetup.res4a_branch1.Bias,"Weights",trainingSetup.res4a_branch1.Weights)

    batchNormalizationLayer("Name","bn4a_branch1","Offset",trainingSetup.bn4a_branch1.Offset,"S
cale",trainingSetup.bn4a_branch1.Scale,"TrainedMean",trainingSetup.bn4a_branch1.TrainedMea
n,"TrainedVariance",trainingSetup.bn4a_branch1.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    additionLayer(2,"Name","res4a")
    reluLayer("Name","res4a_relu");
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([3
3],256,"Name","res4b_branch2a","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res4b_branch2a.Bias,"Weights",trainingSetup.res4b_branch2a.Weights)

    batchNormalizationLayer("Name","bn4b_branch2a","Offset",trainingSetup.bn4b_branch2a.Offset,
"Scale",trainingSetup.bn4b_branch2a.Scale,"TrainedMean",trainingSetup.bn4b_branch2a.Trained
Mean,"TrainedVariance",trainingSetup.bn4b_branch2a.TrainedVariance)
    reluLayer("Name","res4b_branch2a_relu")
    convolution2dLayer([3
3],256,"Name","res4b_branch2b","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res4b_branch2b.Bias,"Weights",trainingSetup.res4b_branch2b.Weights)

    batchNormalizationLayer("Name","bn4b_branch2b","Offset",trainingSetup.bn4b_branch2b.Offset
,"Scale",trainingSetup.bn4b_branch2b.Scale,"TrainedMean",trainingSetup.bn4b_branch2b.Trained
Mean,"TrainedVariance",trainingSetup.bn4b_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    additionLayer(2,"Name","res4b")
    reluLayer("Name","res4b_relu");
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([3
3],512,"Name","res5a_branch2a","BiasLearnRateFactor",0,"Padding",[1 1 1 1],"Stride",[2
2],"Bias",trainingSetup.res5a_branch2a.Bias,"Weights",trainingSetup.res5a_branch2a.Weights)

```

---

---

```

batchNormalizationLayer("Name","bn5a_branch2a","Offset",trainingSetup.bn5a_branch2a.Offset,
"Scale",trainingSetup.bn5a_branch2a.Scale,"TrainedMean",trainingSetup.bn5a_branch2a.Trained
Mean,"TrainedVariance",trainingSetup.bn5a_branch2a.TrainedVariance)
    reluLayer("Name","res5a_branch2a_relu")
    convolution2dLayer([3
3],512,"Name","res5a_branch2b","BiasLearnRateFactor",0,"Padding",[1          1          1
1],"Bias",trainingSetup.res5a_branch2b.Bias,"Weights",trainingSetup.res5a_branch2b.Weights)

batchNormalizationLayer("Name","bn5a_branch2b","Offset",trainingSetup.bn5a_branch2b.Offset,
"Scale",trainingSetup.bn5a_branch2b.Scale,"TrainedMean",trainingSetup.bn5a_branch2b.Trained
Mean,"TrainedVariance",trainingSetup.bn5a_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([1          1],512,"Name","res5a_branch1","BiasLearnRateFactor",0,"Stride",[2
2],"Bias",trainingSetup.res5a_branch1.Bias,"Weights",trainingSetup.res5a_branch1.Weights)

batchNormalizationLayer("Name","bn5a_branch1","Offset",trainingSetup.bn5a_branch1.Offset,"S
cale",trainingSetup.bn5a_branch1.Scale,"TrainedMean",trainingSetup.bn5a_branch1.TrainedMea
n,"TrainedVariance",trainingSetup.bn5a_branch1.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    additionLayer(2,"Name","res5a")
    reluLayer("Name","res5a_relu")];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    convolution2dLayer([3
3],512,"Name","res5b_branch2a","BiasLearnRateFactor",0,"Padding",[1          1          1
1],"Bias",trainingSetup.res5b_branch2a.Bias,"Weights",trainingSetup.res5b_branch2a.Weights)

batchNormalizationLayer("Name","bn5b_branch2a","Offset",trainingSetup.bn5b_branch2a.Offset,
"Scale",trainingSetup.bn5b_branch2a.Scale,"TrainedMean",trainingSetup.bn5b_branch2a.Trained
Mean,"TrainedVariance",trainingSetup.bn5b_branch2a.TrainedVariance)
    reluLayer("Name","res5b_branch2a_relu")
    convolution2dLayer([3
3],512,"Name","res5b_branch2b","BiasLearnRateFactor",0,"Padding",[1          1          1
1],"Bias",trainingSetup.res5b_branch2b.Bias,"Weights",trainingSetup.res5b_branch2b.Weights)

batchNormalizationLayer("Name","bn5b_branch2b","Offset",trainingSetup.bn5b_branch2b.Offset
,"Scale",trainingSetup.bn5b_branch2b.Scale,"TrainedMean",trainingSetup.bn5b_branch2b.Trained
Mean,"TrainedVariance",trainingSetup.bn5b_branch2b.TrainedVariance)];
lgraph = addLayers(lgraph,tempLayers);

tempLayers = [
    additionLayer(2,"Name","res5b")
    reluLayer("Name","res5b_relu")
    globalAveragePooling2dLayer("Name","pool5")
    fullyConnectedLayer(3,"Name","fc")
    softmaxLayer("Name","prob")
    classificationLayer("Name","classoutput")];
lgraph = addLayers(lgraph,tempLayers);

Ⓒ Clean up helper variable:
clear tempLayers;

Ⓓ Connect layer branches:
lgraph = connectLayers(lgraph,"pool1","res2a_branch2a");
lgraph = connectLayers(lgraph,"pool1","res2a/in2");
lgraph = connectLayers(lgraph,"bn2a_branch2b","res2a/in1");
lgraph = connectLayers(lgraph,"res2a_relu","res2b_branch2a");
lgraph = connectLayers(lgraph,"res2a_relu","res2b/in2");
lgraph = connectLayers(lgraph,"bn2b_branch2b","res2b/in1");

```

---

---

```
lgraph = connectLayers(lgraph,"res2b_relu","res3a_branch2a");
lgraph = connectLayers(lgraph,"res2b_relu","res3a_branch1");
lgraph = connectLayers(lgraph,"bn3a_branch2b","res3a/in1");
lgraph = connectLayers(lgraph,"bn3a_branch1","res3a/in2");
lgraph = connectLayers(lgraph,"res3a_relu","res3b_branch2a");
lgraph = connectLayers(lgraph,"res3a_relu","res3b/in2");
lgraph = connectLayers(lgraph,"bn3b_branch2b","res3b/in1");
lgraph = connectLayers(lgraph,"res3b_relu","res4a_branch2a");
lgraph = connectLayers(lgraph,"res3b_relu","res4a_branch1");
lgraph = connectLayers(lgraph,"bn4a_branch2b","res4a/in1");
lgraph = connectLayers(lgraph,"bn4a_branch1","res4a/in2");
lgraph = connectLayers(lgraph,"res4a_relu","res4b_branch2a");
lgraph = connectLayers(lgraph,"res4a_relu","res4b/in2");
lgraph = connectLayers(lgraph,"bn4b_branch2b","res4b/in1");
lgraph = connectLayers(lgraph,"res4b_relu","res5a_branch2a");
lgraph = connectLayers(lgraph,"res4b_relu","res5a_branch1");
lgraph = connectLayers(lgraph,"bn5a_branch2b","res5a/in1");
lgraph = connectLayers(lgraph,"bn5a_branch1","res5a/in2");
lgraph = connectLayers(lgraph,"res5a_relu","res5b_branch2a");
lgraph = connectLayers(lgraph,"res5a_relu","res5b/in2");
lgraph = connectLayers(lgraph,"bn5b_branch2b","res5b/in1");
```

---

#### ① Train network:

```
[net, traininfo] = trainNetwork(augimdsTrain,lgraph,opts);
```

---

This code was used in the MATLAB release R2023b.

## References

1. Hodson, R. Inflammatory bowel disease. *Nature* **2016**, 540, S97. <https://doi.org/10.1038/540S97a>.
2. Sairenji, T.; Collins, K.L.; Evans, D.V. An Update on Inflammatory Bowel Disease. *Prim Care* **2017**, 44, 673-692. <https://doi.org/10.1016/j.pop.2017.07.010>.
3. Bruner, L.P.; White, A.M.; Proksell, S. Inflammatory Bowel Disease. *Prim Care* **2023**, 50, 411-427. <https://doi.org/10.1016/j.pop.2023.03.009>.
4. Din, S.; Wong, K.; Mueller, M.F.; Oniscu, A.; Hewinson, J.; Black, C.J.; Miller, M.L.; Jimenez-Sanchez, A.; Rabbie, R.; Rashid, M.; et al. Mutational Analysis Identifies Therapeutic Biomarkers in Inflammatory Bowel Disease-Associated Colorectal Cancers. *Clin Cancer Res* **2018**, 24, 5133-5142. <https://doi.org/10.1158/1078-0432.CCR-17-3713>.
5. Halliday, G.; Porter, R.J.; Black, C.J.; Arends, M.J.; Din, S. c-MET immunohistochemical expression in sporadic and inflammatory bowel disease associated lesions. *World J Gastroenterol* **2022**, 28, 1338-1346. <https://doi.org/10.3748/wjg.v28.i13.1338>.
6. Hemmer, A.; Forest, K.; Rath, J.; Bowman, J. Inflammatory Bowel Disease: A Concise Review. *S D Med* **2023**, 76, 416-423.
7. Khor, B.; Gardet, A.; Xavier, R.J. Genetics and pathogenesis of inflammatory bowel disease. *Nature* **2011**, 474, 307-317. <https://doi.org/10.1038/nature10209>.
8. Porter, R.J.; Arends, M.J.; Churchhouse, A.M.D.; Din, S. Inflammatory Bowel Disease-Associated Colorectal Cancer: Translational Risks from Mechanisms to Medicines. *J Crohns Colitis* **2021**, 15, 2131-2141. <https://doi.org/10.1093/ecco-jcc/jjab102>.
9. Zhang, Y.Z.; Li, Y.Y. Inflammatory bowel disease: pathogenesis. *World J Gastroenterol* **2014**, 20, 91-99. <https://doi.org/10.3748/wjg.v20.i1.91>.
10. Singh, N.; Bernstein, C.N. Environmental risk factors for inflammatory bowel disease. *United European Gastroenterol J* **2022**, 10, 1047-1053. <https://doi.org/10.1002/ueg2.12319>.
11. Qiu, P.; Ishimoto, T.; Fu, L.; Zhang, J.; Zhang, Z.; Liu, Y. The Gut Microbiota in Inflammatory Bowel Disease. *Front Cell Infect Microbiol* **2022**, 12, 733992. <https://doi.org/10.3389/fcimb.2022.733992>.



12. Jarmakiewicz-Czaja, S.; Zielinska, M.; Sokal, A.; Filip, R. Genetic and Epigenetic Etiology of Inflammatory Bowel Disease: An Update. *Genes (Basel)* **2022**, *13*. <https://doi.org/10.3390/genes13122388>.
13. Graham, D.B.; Xavier, R.J. Pathway paradigms revealed from the genetics of inflammatory bowel disease. *Nature* **2020**, *578*, 527-539. <https://doi.org/10.1038/s41586-020-2025-2>.
14. Saez, A.; Gomez-Bris, R.; Herrero-Fernandez, B.; Mingorance, C.; Rius, C.; Gonzalez-Granado, J.M. Innate Lymphoid Cells in Intestinal Homeostasis and Inflammatory Bowel Disease. *Int J Mol Sci* **2021**, *22*. <https://doi.org/10.3390/ijms22147618>.
15. Saez, A.; Herrero-Fernandez, B.; Gomez-Bris, R.; Sanchez-Martinez, H.; Gonzalez-Granado, J.M. Pathophysiology of Inflammatory Bowel Disease: Innate Immune System. *Int J Mol Sci* **2023**, *24*. <https://doi.org/10.3390/ijms24021526>.
16. Lu, Q.; Yang, M.F.; Liang, Y.J.; Xu, J.; Xu, H.M.; Nie, Y.Q.; Wang, L.S.; Yao, J.; Li, D.F. Immunology of Inflammatory Bowel Disease: Molecular Mechanisms and Therapeutics. *J Inflamm Res* **2022**, *15*, 1825-1844. <https://doi.org/10.2147/JIR.S353038>.
17. Kaplan, G.G.; Windsor, J.W. The four epidemiological stages in the global evolution of inflammatory bowel disease. *Nat Rev Gastroenterol Hepatol* **2021**, *18*, 56-66. <https://doi.org/10.1038/s41575-020-00360-x>.
18. Agrawal, M.; Jess, T. Implications of the changing epidemiology of inflammatory bowel disease in a changing world. *United European Gastroenterol J* **2022**, *10*, 1113-1120. <https://doi.org/10.1002/ueg2.12317>.
19. Narula, N.; Wong, E.C.L.; Dehghan, M.; Mente, A.; Rangarajan, S.; Lanas, F.; Lopez-Jaramillo, P.; Rohatgi, P.; Lakshmi, P.V.M.; Varma, R.P.; et al. Association of ultra-processed food intake with risk of inflammatory bowel disease: prospective cohort study. *BMJ* **2021**, *374*, n1554. <https://doi.org/10.1136/bmj.n1554>.
20. Burisch, J.; Zhao, M.; Odes, S.; De Cruz, P.; Vermeire, S.; Bernstein, C.N.; Kaplan, G.G.; Duricova, D.; Greenberg, D.; Melberg, H.O.; et al. The cost of inflammatory bowel disease in high-income settings: a Lancet Gastroenterology & Hepatology Commission. *Lancet Gastroenterol Hepatol* **2023**, *8*, 458-492. [https://doi.org/10.1016/S2468-1253\(23\)00003-1](https://doi.org/10.1016/S2468-1253(23)00003-1).
21. Buie, M.J.; Quan, J.; Windsor, J.W.; Coward, S.; Hansen, T.M.; King, J.A.; Kotze, P.G.; Gearry, R.B.; Ng, S.C.; Mak, J.W.Y.; et al. Global Hospitalization Trends for Crohn's Disease and Ulcerative Colitis in the 21st Century: A Systematic Review With Temporal Analyses. *Clin Gastroenterol Hepatol* **2023**, *21*, 2211-2221. <https://doi.org/10.1016/j.cgh.2022.06.030>.
22. Wijnands, A.M.; Elias, S.G.; Dekker, E.; Fidder, H.H.; Hoentjen, F.; Ten Hove, J.R.; Maljaars, P.W.J.; van der Meulen-de Jong, A.E.; Mooiweer, E.; Ouwehand, R.J.; et al. Smoking and colorectal neoplasia in patients with inflammatory bowel disease: Dose-effect relationship. *United European Gastroenterol J* **2023**, *11*, 612-620. <https://doi.org/10.1002/ueg2.12426>.
23. Ham, N.S.; Hwang, S.W.; Oh, E.H.; Kim, J.; Lee, H.S.; Park, S.H.; Yang, D.H.; Ye, B.D.; Byeon, J.S.; Myung, S.J.; et al. Influence of Severe Vitamin D Deficiency on the Clinical Course of Inflammatory Bowel Disease. *Dig Dis Sci* **2021**, *66*, 587-596. <https://doi.org/10.1007/s10620-020-06207-4>.
24. Mark A Peppercorn, Adam S Cheifetz. Definitions, epidemiology, and risk factors for inflammatory bowel disease. In: UpToDate, Sunanda V Kane (Ed), Wolters Kluwer. (Accessed on December 11, 2024.).
25. Ungaro, R.; Mehandru, S.; Allen, P.B.; Peyrin-Biroulet, L.; Colombel, J.F. Ulcerative colitis. *Lancet* **2017**, *389*, 1756-1770. [https://doi.org/10.1016/S0140-6736\(16\)32126-2](https://doi.org/10.1016/S0140-6736(16)32126-2).
26. Gomez-Bris, R.; Saez, A.; Herrero-Fernandez, B.; Rius, C.; Sanchez-Martinez, H.; Gonzalez-Granado, J.M. CD4 T-Cell Subsets and the Pathophysiology of Inflammatory Bowel Disease. *Int J Mol Sci* **2023**, *24*. <https://doi.org/10.3390/ijms24032696>.

27. Cui, G.; Yuan, A.; Sorbye, S.W.; Florholmen, J. Th9 and Th17 Cells in Human Ulcerative Colitis-Associated Dysplastic Lesions. *Clin Med Insights Oncol* **2024**, *18*, 11795549241301358. <https://doi.org/10.1177/11795549241301358>.
28. Gerlach, K.; Lechner, K.; Popp, V.; Offensperger, L.; Zundler, S.; Wiendl, M.; Becker, E.; Atreya, R.; Rath, T.; Neurath, M.F.; et al. The JAK1/3 inhibitor tofacitinib suppresses T cell homing and activation in chronic intestinal inflammation. *J Crohns Colitis* **2020**. <https://doi.org/10.1093/ecco-jcc/jjaa162>.
29. Mitsialis, V.; Wall, S.; Liu, P.; Ordovas-Montanes, J.; Parmet, T.; Vukovic, M.; Spencer, D.; Field, M.; McCourt, C.; Toothaker, J.; et al. Single-Cell Analyses of Colon and Blood Reveal Distinct Immune Cell Signatures of Ulcerative Colitis and Crohn's Disease. *Gastroenterology* **2020**, *159*, 591-608 e510. <https://doi.org/10.1053/j.gastro.2020.04.074>.
30. Shmuel-Galia, L.; Humphries, F.; Lei, X.; Ceglia, S.; Wilson, R.; Jiang, Z.; Ketelut-Carneiro, N.; Foley, S.E.; Pechhold, S.; Houghton, J.; et al. Dysbiosis exacerbates colitis by promoting ubiquitination and accumulation of the innate immune adaptor STING in myeloid cells. *Immunity* **2021**, *54*, 1137-1153 e1138. <https://doi.org/10.1016/j.immuni.2021.05.008>.
31. DeGruttola, A.K.; Low, D.; Mizoguchi, A.; Mizoguchi, E. Current Understanding of Dysbiosis in Disease in Human and Animal Models. *Inflamm Bowel Dis* **2016**, *22*, 1137-1150. <https://doi.org/10.1097/MIB.0000000000000750>.
32. Schroeder, K.W.; Tremaine, W.J.; Ilstrup, D.M. Coated oral 5-aminosalicylic acid therapy for mildly to moderately active ulcerative colitis. A randomized study. *N Engl J Med* **1987**, *317*, 1625-1629. <https://doi.org/10.1056/NEJM198712243172603>.
33. Silverberg, M.S.; Satsangi, J.; Ahmad, T.; Arnott, I.D.; Bernstein, C.N.; Brant, S.R.; Caprilli, R.; Colombel, J.F.; Gasche, C.; Geboes, K.; et al. Toward an integrated clinical, molecular and serological classification of inflammatory bowel disease: report of a Working Party of the 2005 Montreal World Congress of Gastroenterology. *Can J Gastroenterol* **2005**, *19 Suppl A*, 5A-36A. <https://doi.org/10.1155/2005/269076>.
34. Truelove, S.C.; Witts, L.J. Cortisone in ulcerative colitis; final report on a therapeutic trial. *Br Med J* **1955**, *2*, 1041-1048. <https://doi.org/10.1136/bmj.2.4947.1041>.
35. Geboes, K.; Riddell, R.; Ost, A.; Jensfelt, B.; Persson, T.; Lofberg, R. A reproducible grading scale for histological assessment of inflammation in ulcerative colitis. *Gut* **2000**, *47*, 404-409. <https://doi.org/10.1136/gut.47.3.404>.
36. Fabian, O.; Kamaradova, K. Morphology of inflammatory bowel diseases (IBD). *Cesk Patol* **2022**, *58*, 27-37.
37. Feakins, R.M. Ulcerative colitis or Crohn's disease? Pitfalls and problems. *Histopathology* **2014**, *64*, 317-335. <https://doi.org/10.1111/his.12263>.
38. El-Zimaity, H.; Shaffer, S.R.; Riddell, R.H.; Pai, R.K.; Bernstein, C.N. Beyond Neutrophils for Predicting Relapse and Remission in Ulcerative Colitis. *J Crohns Colitis* **2023**, *17*, 767-776. <https://doi.org/10.1093/ecco-jcc/jjac178>.
39. Gupta, R.B.; Harpaz, N.; Itzkowitz, S.; Hossain, S.; Matula, S.; Kornbluth, A.; Bodian, C.; Ullman, T. Histologic inflammation is a risk factor for progression to colorectal neoplasia in ulcerative colitis: a cohort study. *Gastroenterology* **2007**, *133*, 1099-1105; quiz 1340-1091. <https://doi.org/10.1053/j.gastro.2007.08.001>.
40. Gros, B.; Kaplan, G.G. Ulcerative Colitis in Adults: A Review. *JAMA* **2023**, *330*, 951-965. <https://doi.org/10.1001/jama.2023.15389>.
41. Russell D Cohen, Adam C Stein. Management of moderate to severe ulcerative colitis in adults. In: UpToDate, Sunanda V Kane (Ed), Wolters Kluwer. (Accessed on December, 2024.).
42. Jana Al Hashash, Miguel Regueiro. Medical management of low-risk adult patients with mild to moderate ulcerative colitis. In: UpToDate, Kristen M Robson (Ed), Wolters Kluwer. (Accessed on December 11, 2024.).

43. Le Berre, C.; Honap, S.; Peyrin-Biroulet, L. Ulcerative colitis. *Lancet* **2023**, *402*, 571-584. [https://doi.org/10.1016/S0140-6736\(23\)00966-2](https://doi.org/10.1016/S0140-6736(23)00966-2).
44. Marshall, D.A.; MacDonald, K.V.; Kao, D.; Bernstein, C.N.; Kaplan, G.G.; Jijon, H.; Hazlewood, G.; Panaccione, R.; Nasser, Y.; Raman, M.; et al. Patient preferences for active ulcerative colitis treatments and fecal microbiota transplantation. *Ther Adv Chronic Dis* **2024**, *15*, 20406223241239168. <https://doi.org/10.1177/20406223241239168>.
45. Global Cancer Observatory. International Agency for Research on Cancer. World Health Organization. Available at: <https://gco.iarc.fr/> (Accessed on December 13, 2023). Available online: (accessed on
46. Finlay A Macrae. Epidemiology and risk factors for colorectal cancer. In: UpToDate, Richard M Goldberg, David Seres (Eds), Wolters Kluwer. (Accessed on December 12, 2024.).
47. Baidoun, F.; Elshiw, K.; Elkerai, Y.; Merjane, Z.; Khoudari, G.; Sarmini, M.T.; Gad, M.; Al-Husseini, M.; Saad, A. Colorectal Cancer Epidemiology: Recent Trends and Impact on Outcomes. *Curr Drug Targets* **2021**, *22*, 998-1009. <https://doi.org/10.2174/1389450121999201117115717>.
48. Dekker, E.; Tanis, P.J.; Vleugels, J.L.A.; Kasi, P.M.; Wallace, M.B. Colorectal cancer. *Lancet* **2019**, *394*, 1467-1480. [https://doi.org/10.1016/S0140-6736\(19\)32319-0](https://doi.org/10.1016/S0140-6736(19)32319-0).
49. Patel, S.G.; Karlitz, J.J.; Yen, T.; Lieu, C.H.; Boland, C.R. The rising tide of early-onset colorectal cancer: a comprehensive review of epidemiology, clinical features, biology, risk factors, prevention, and early detection. *Lancet Gastroenterol Hepatol* **2022**, *7*, 262-274. [https://doi.org/10.1016/S2468-1253\(21\)00426-X](https://doi.org/10.1016/S2468-1253(21)00426-X).
50. Sullivan, B.A.; Noujaim, M.; Roper, J. Cause, Epidemiology, and Histology of Polyps and Pathways to Colorectal Cancer. *Gastrointest Endosc Clin N Am* **2022**, *32*, 177-194. <https://doi.org/10.1016/j.giec.2021.12.001>.
51. Yokoyama, S.; Watanabe, T.; Fujita, Y.; Matsumura, S.; Ueda, K.; Nagano, S.; Kinoshita, I.; Murakami, D.; Tabata, H.; Tsuji, T.; et al. Histology of metastatic colorectal cancer in a lymph node. *PLoS One* **2023**, *18*, e0284536. <https://doi.org/10.1371/journal.pone.0284536>.
52. Nagtegaal, I.D.; Hugen, N. The Increasing Relevance of Tumour Histology in Determining Oncological Outcomes in Colorectal Cancer. *Curr Colorectal Cancer Rep* **2015**, *11*, 259-266. <https://doi.org/10.1007/s11888-015-0280-7>.
53. Gulsoy, T.; Baykal Kablan, E. FocalNeXt: A ConvNeXt augmented FocalNet architecture for lung cancer classification from CT-scan images. *Expert Systems with Applications* **2025**, *261*. <https://doi.org/10.1016/j.eswa.2024.125553>.
54. Taatjes, D.J.; Bouffard, N.A.; Barrow, T.; Devitt, K.A.; Gardner, J.A.; Braet, F. Quantitative pixel intensity- and color-based image analysis on minimally compressed files: implications for whole-slide imaging. *Histochem Cell Biol* **2019**, *152*, 13-23. <https://doi.org/10.1007/s00418-019-01783-7>.
55. Hofener, H.; Homeyer, A.; Weiss, N.; Molin, J.; Lundstrom, C.F.; Hahn, H.K. Deep learning nuclei detection: A simple approach can deliver state-of-the-art results. *Comput Med Imaging Graph* **2018**, *70*, 43-52. <https://doi.org/10.1016/j.compmedimag.2018.08.010>.
56. MathWorks. MATLAB for Artificial Intelligence. Design AI models and AI-driven systems. Website: <https://www.mathworks.com/> (Accessed on December 12, 2024). Available online: (accessed on
57. Ewaeed, N.A.; Abed, H.N.; Abed, S.N. Detecting and Classifying Household Insects in Iraq by using Transfer Learning Models. *Journal of Advanced Research in Applied Sciences and Engineering Technology* **2025**, *50*, 21-33. <https://doi.org/10.37934/araset.50.1.2133>.
58. Xu, H.; Usuyama, N.; Bagga, J.; Zhang, S.; Rao, R.; Naumann, T.; Wong, C.; Gero, Z.; Gonzalez, J.; Gu, Y.; et al. A whole-slide foundation model for digital pathology from real-world data. *Nature* **2024**, *630*, 181-188. <https://doi.org/10.1038/s41586-024-07441-w>.



59. Das, N.; Das, S. Attention-UNet architectures with pretrained backbones for multi-class cardiac MR image segmentation. *Curr Probl Cardiol* **2024**, *49*, 102129. <https://doi.org/10.1016/j.cpcardiol.2023.102129>.
60. Jiang, X.; Hu, Z.; Wang, S.; Zhang, Y. Deep Learning for Medical Image-Based Cancer Diagnosis. *Cancers (Basel)* **2023**, *15*. <https://doi.org/10.3390/cancers15143608>.
61. Weller, J.H.; Scheese, D.; Tragesser, C.; Yi, P.H.; Alaish, S.M.; Hackam, D.J. Artificial Intelligence vs. Doctors: Diagnosing Necrotizing Enterocolitis on Abdominal Radiographs. *J Pediatr Surg* **2024**, *59*, 161592. <https://doi.org/10.1016/j.jpedsurg.2024.06.001>.
62. Khan, H.A.; Jue, W.; Mushtaq, M.; Mushtaq, M.U. Brain tumor classification in MRI image using convolutional neural network. *Math Biosci Eng* **2020**, *17*, 6203-6216. <https://doi.org/10.3934/mbe.2020328>.
63. Karimi, D.; Dou, H.; Gholipour, A. Medical Image Segmentation Using Transformer Networks. *IEEE Access* **2022**, *10*, 29322-29332. <https://doi.org/10.1109/access.2022.3156894>.
64. Wang, X.; Yang, S.; Zhang, J.; Wang, M.; Zhang, J.; Yang, W.; Huang, J.; Han, X. Transformer-based unsupervised contrastive learning for histopathological image classification. *Med Image Anal* **2022**, *81*, 102559. <https://doi.org/10.1016/j.media.2022.102559>.
65. Lama, N.; Kasmi, R.; Hagerty, J.R.; Stanley, R.J.; Young, R.; Miinch, J.; Nepal, J.; Nambisan, A.; Stoecker, W.V. ChimeraNet: U-Net for Hair Detection in Dermoscopic Skin Lesion Images. *J Digit Imaging* **2023**, *36*, 526-535. <https://doi.org/10.1007/s10278-022-00740-6>.
66. TensorFlow. Website: <https://www.tensorflow.org/> (Accessed on December 12, 2024). Available online: (accessed on
67. TensorFlow Keras Basic image classification. Website: <https://www.tensorflow.org/tutorials/keras/classification> (Accessed on December 12, 2024). Available online: (accessed on
68. PyTorch, get started. Website: <https://pytorch.org/> (Accessed on December 12, 2024). Available online: (accessed on
69. ONNX, open neural network exchange. Website: <https://onnx.ai/> (Accessed on December 12, 2024). Available online: (accessed on
70. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. Deep Residual Learning for Image Recognition. arXiv:1512.03385v1.
71. UniProt, C. UniProt: the Universal Protein Knowledgebase in 2023. *Nucleic Acids Res* **2023**, *51*, D523-D531. <https://doi.org/10.1093/nar/gkac1052>.
72. Peng, D.H.; Rodriguez, B.L.; Diao, L.; Chen, L.; Wang, J.; Byers, L.A.; Wei, Y.; Chapman, H.A.; Yamauchi, M.; Behrens, C.; et al. Collagen promotes anti-PD-1/PD-L1 resistance in cancer through LAIR1-dependent CD8(+) T cell exhaustion. *Nat Commun* **2020**, *11*, 4520. <https://doi.org/10.1038/s41467-020-18298-8>.
73. Van Laethem, F.; Donaty, L.; Tchernonog, E.; Lacheretz-Szablewski, V.; Russello, J.; Buthiau, D.; Almeras, M.; Moreaux, J.; Bret, C. LAIR1, an ITIM-Containing Receptor Involved in Immune Disorders and in Hematological Neoplasms. *Int J Mol Sci* **2022**, *23*. <https://doi.org/10.3390/ijms232416136>.
74. Xu, W.; Zhao, X.; Wang, X.; Feng, H.; Gou, M.; Jin, W.; Wang, X.; Liu, X.; Dong, C. The Transcription Factor Tox2 Drives T Follicular Helper Cell Development via Regulating Chromatin Accessibility. *Immunity* **2019**, *51*, 826-839 e825. <https://doi.org/10.1016/j.immuni.2019.10.006>.
75. Carreras, J.; Roncador, G.; Hamoudi, R. Ulcerative Colitis, LAIR1 and TOX2 expression and Colorectal Cancer Deep Learning Image Classification Using Convolutional Neural Networks. Preprints 2024, 2024110211. <https://doi.org/10.20944/preprints202411.0211.v1>.
76. Carreras, J.; Kikuti, Y.Y.; Bea, S.; Miyaoka, M.; Hiraiwa, S.; Ikoma, H.; Nagao, R.; Tomita, S.; Martin-Garcia, D.; Salaverria, I.; et al. Clinicopathological characteristics and genomic profile of primary sinonasal tract

- diffuse large B cell lymphoma (DLBCL) reveals gain at 1q31 and RGS1 encoding protein; high RGS1 immunohistochemical expression associates with poor overall survival in DLBCL not otherwise specified (NOS). *Histopathology* **2017**, *70*, 595-621. <https://doi.org/10.1111/his.13106>.
77. Carreras, J.; Kikuti, Y.Y.; Hiraiwa, S.; Miyaoka, M.; Tomita, S.; Ikoma, H.; Ito, A.; Kondo, Y.; Itoh, J.; Roncador, G.; et al. High PTX3 expression is associated with a poor prognosis in diffuse large B-cell lymphoma. *Cancer Sci* **2022**, *113*, 334-348. <https://doi.org/10.1111/cas.15179>.
  78. Carreras, J.; Yukie Kikuti, Y.; Miyaoka, M.; Hiraiwa, S.; Tomita, S.; Ikoma, H.; Kondo, Y.; Shiraiwa, S.; Ando, K.; Sato, S.; et al. Genomic Profile and Pathologic Features of Diffuse Large B-Cell Lymphoma Subtype of Methotrexate-associated Lymphoproliferative Disorder in Rheumatoid Arthritis Patients. *Am J Surg Pathol* **2018**, *42*, 936-950. <https://doi.org/10.1097/PAS.0000000000001071>.
  79. Carreras, J.; Kikuti, Y.Y.; Roncador, G.; Miyaoka, M.; Hiraiwa, S.; Tomita, S.; Ikoma, H.; Kondo, Y.; Ito, A.; Shiraiwa, S.; et al. High Expression of Caspase-8 Associated with Improved Survival in Diffuse Large B-Cell Lymphoma: Machine Learning and Artificial Neural Networks Analyses. *BioMedInformatics* **2021**, *1*, 18-46. <https://doi.org/10.3390/biomedinformatics1010003>.
  80. Carreras, J. Celiac Disease Deep Learning Image Classification Using Convolutional Neural Networks. *J. Imaging* **2024**, *10*, 200. <https://doi.org/10.3390/jimaging10080200>.
  81. MathWorks. Batch normalization layer. Available online: <https://www.mathworks.com/help/deeplearning/ref/nnet.cnn.layer.batchnormalizationlayer.html> (accessed on 28 July 2024).
  82. Carreras, J.; Kikuti, Y.Y.; Miyaoka, M.; Roncador, G.; Garcia, J.F.; Hiraiwa, S.; Tomita, S.; Ikoma, H.; Kondo, Y.; Ito, A.; et al. Integrative Statistics, Machine Learning and Artificial Intelligence Neural Network Analysis Correlated CSF1R with the Prognosis of Diffuse Large B-Cell Lymphoma. *Hemato* **2021**, *2*, 182-206. <https://doi.org/10.3390/hemato2020011>.
  83. Carreras, J.; Nakamura, N. Artificial Intelligence, Lymphoid Neoplasms, and Prediction of MYC, BCL2, and BCL6 Gene Expression Using a Pan-Cancer Panel in Diffuse Large B-Cell Lymphoma. *Hemato* **2024**, *5*, 119-143. <https://doi.org/10.3390/hemato5020011>.
  84. Carreras, J.; Yukie Kikuti, Y.; Miyaoka, M.; Miyahara, S.; Roncador, G.; Hamoudi, R.; Nakamura, N. Artificial Intelligence Analysis and Reverse Engineering of Molecular Subtypes of Diffuse Large B-Cell Lymphoma Using Gene Expression Data. *BioMedInformatics* **2024**, *4*, 295-320. <https://doi.org/10.3390/biomedinformatics4010017>