HAPPY: A deep learning pipeline for mapping cell-to-tissue graphs across placenta histology whole slide images

Claudia Vanea, Jelisaveta Džigurski, Valentina Rukins, Omri Dodi, Siim Siigur, Liis Salumäe, Karen Meir, W. Tony Parks, Drorith Hochner-Celnikier, Abigail Fraser, Hagit Hochner, Triin Laisk, Linda M. Ernst, Cecilia M. Lindgren, Christoffer Nellåker

Author information

These authors contributed equally: Christoffer Nellåker and Cecilia M. Lindgren

Authors and Affiliations

Nuffield Department of Women's & Reproductive Health, University of Oxford
Claudia Vanea, Christoffer Nellåker

Big Data Institute, University of Oxford
Claudia Vanea, Cecilia M. Lindgren, Christoffer Nellåker

Institute of Genomics, University of Tartu
Jelisaveta Džigurski, Valentina Rukins, Triin Laisk

Faculty of Medicine, Hadassah Hebrew University Medical Center
Omri Dodi, Karen Meir, Drorith Hochner-Celnikier

Department of Pathology, Tartu University Hospital
Siim Siigur, Liis Salumäe

Department of Laboratory Medicine & Pathobiology, University of Toronto
W. Tony Parks

Population Health Sciences, Bristol Medical School, University of Bristol
Abigail Fraser

Braun School of Public Health, Hebrew University of Jerusalem
Hagit Hochner

Northshore University HealthSystem, University of Chicago
Linda M. Ernst

Corresponding author

Correspondence to Christoffer Nellåker (christoffer.nellaker@bdi.ox.ac.uk) or Claudia Vanea (claudia.vanea@wrh.ox.ac.uk)

Abstract

Accurate placenta pathology assessment is essential for managing maternal and newborn health, but the placenta's heterogeneity and temporal variability pose challenges for histology analysis. To address this issue, we developed the 'Histology Analysis Pipeline.PY' (HAPPY), a deep learning hierarchical method for quantifying the variability of cells and micro-anatomical tissue structures across placenta histology whole slide images. HAPPY differs from patch-based features or segmentation approaches by following an interpretable biological hierarchy, representing cells and cellular communities within tissues at a single-cell resolution across whole slide images. We present a set of quantitative metrics from healthy term placentas as a baseline for future assessments of placenta health and we show how these metrics deviate in placentas with clinically significant placental infarction. HAPPY's cell and tissue predictions closely replicate those from independent clinical experts and placental biology literature.

Introduction

Accurate placenta pathology assessment is essential for clinical management of mother and newborn health[1–3]. Despite this, the placenta remains an understudied and poorly understood organ[1,4–13]. It is the first organ formed by the developing fetus and is heterogeneous[14–16] and rapidly evolving[2–4,17]. The placenta's high spatial and temporal variability presents a challenge for robust, reproducible histological analysis and detection of pathological changes. The difficulty of this task is reflected in low interobserver reliability among pathologists for many diagnostically-relevant placental features (i.e. gestational age[18], villus maturity[19], maternal vascular malperfusion[20])[2,21,22]. Placenta histology slides commonly contain upwards of a million cells comprising tens of thousands of tissue microstructures. High-throughput, quantitative and objective metrics of placental biology are therefore valuable for placental investigations in both clinical and research settings.

Digital pathology has the potential to provide these high-throughput metrics, allowing for automated processing of histology slides at scale. In recent years, deep learning is becoming a gold standard for digital pathology, with success in cancer survival analysis[23–25] and tumour micro-environment modelling[26–30]. However, there has been little application of these approaches to placenta histology[16,31,32], where reliable quantification of micro-anatomical structures across slides is vital to determining key features of health assessment such as placental maturity[3,15,17,33].

Established deep learning approaches for quantifying micro-anatomical structures in histology have been largely based on tissue segmentation[34–38] or prediction at the patch level[30,39–41]. Tissue segmentation typically requires a large amount of precise manually curated annotations to perform well[35–38] and while patch-based approaches only require patch-level labels, prediction resolution is limited to patch size. Additionally, as both approaches operate within fixed patches, they are unable to utilise contextual information external to a patch and are vulnerable to changes in patch construction[42].

In recent years, the representation of cells as a spatial graph[24,43–53] has emerged as a promising way to discover new cellular phenotypes[45,49,53], make slide-level

predictions[24,44,47,52], and hierarchically cluster cells into tissue microstructures[27,46,50,51]. However, for the most common imaging modality, Hematoxylin and Eosin (H&E) stained histology, these have only been applied for patch-level[27,46,48] and slide-level prediction[24,44,47,52] or on graphs restricted to fixed subregions[46,50]. Here we present a pipeline for analysing tissue microstructures at the single cell level by building whole slide spatial cell graphs across H&E histology with dynamic sampling of graph regions and learnt cellular community aggregation. This pipeline leverages the relatively large cell sample size within a slide to overcome uncertainty and variance in any one classification and to achieve robust tissue microstructure classification.

Results

Automated quantification of healthy variability in placenta histology

We present HAPPY (Histology Analysis Pipeline.PY), a novel method for quantifying cells and micro-anatomical tissue structures across Hematoxylin and Eosin (H&E) stained placenta histology whole slide images (WSIs). Our approach is inspired by the biological hierarchy of the organ, from locating all nuclei across a WSI, to classifying their cell types, to identifying the tissue microstructures that those cells comprise. HAPPY can facilitate large-scale morphometric studies of placenta histology, a currently manual, expert-requiring, labour-intensive task.

HAPPY classifies all cells in a placenta parenchyma WSI into one of 11 cell types and 9 tissue microstructure categories. We train nuclei localisation and cell classification models on 11,755 nuclei and 11,789 cells and evaluate on a held-out test set of 2,754 nuclei and 2,381 cells. We train a graph neural network tissue classification model on 468,869 nodes and evaluate on a held-out portion of 179,095 test nodes across microstructures. We compare the tissue classification model's performance with the labels and Cohen's kappa agreement scores of four practising perinatal pathologists over 180 tissues and we use these to validate the ground truth annotations. We show how in five WSIs from five healthy term placentas the predicted proportions of cells, tissues, and composition of tissues correspond to expectations from our current understanding of placental biology. We present these findings as new quantitative metrics for placental health. Finally, we present a pilot case study of clinically significant placental infarction in 12 WSIs from eight term placentas and show how HAPPY identifies biologically-relevant differences in cell and tissue microstructures compared to the healthy group.

HAPPY is structured as a deep learning pipeline in three stages: i) an object detection model for nuclei localisation, ii) an image classification model for cell classification, and iii) a graph neural network (GNN) for tissue classification (Fig. 1). From the nuclei locations and cell predictions, we construct a spatial cell graph across the WSI. To model the interactions of cells within tissue microstructures, we use inductive message-passing node classification across the constructed cell graph. This hierarchical, message-passing approach induces our models to use the constituent cellular information to understand tissue microstructures. By reducing the WSI dimensionality to cell representations we are able to model tissue microstructures at the single cell resolution. The trained tissue model, based on an inductive GNN, is robust to input graphs of any shape or size allowing for application across other

WSIs. This is to our knowledge the first time inductive message-passing graph neural networks have been applied for node classification to cell graphs spanning entire human WSIs. We apply this to generate the first automated whole slide-scale quantification of cells and tissue microstructures in placenta histology.

The HAPPY codebase, training data, and trained models for placenta histology are available at (https://github.com/Nellaker-group/happy). The codebase supports the most commonly used WSI scanner formats[54–56], has additional utilities for creating datasets and visualising outputs, and we provide tutorials for training and inference workflows. HAPPY is presented here applied to the placenta, but by design the codebase can be directly applied to other organ histology, given organ-specific training data. To show that this is valid in principle, we have additionally conducted a preliminary investigation of our nuclei localisation and cell classification models across WSIs of a placenta membrane roll, umbilical cord, a second-trimester placenta with chorioamnionitis, and also across WSIs of other organs in the GTEx dataset (Supplementary Fig. 1).

Evaluation of model performance

We evaluate the deep learning model from each stage of the pipeline on respective unseen held-out test sets (Table 1). The nuclei localisation model achieves a 0.884 F1 score across 2,754 nuclei within 38 images, comparable to F1 scores reported by other nuclei detection models trained for other organs (HoVer-Net achieves an F1 score of 0.756 on the CoNSeP dataset[57] and 0.800 on the PanNuke dataset[58]). The cell classification model, evaluated across 2,381 cells for 11 placental cell types, achieves an overall accuracy of 80% and a top-2 accuracy of 93%, with a 0.969 macro-averaged Receiver Operating Characteristic Area Under Curve (ROC AUC). We show (Fig. 2a) that most misclassifications are within closely related cell differentiation pathways. In Supplementary Table 1, we show how the addition of stain augmentation during training improves model generalisability. See Supplementary Fig. 2 for visualisations of predictions across five WSIs of healthy term placentas.

The graph neural network tissue classification model, evaluated across 149,425 nodes for 9 tissue types, achieves an overall accuracy of 67% and a top-2 and top-3 accuracy of 90% and 96%, with a 0.898 macro-averaged ROC AUC. We show (Fig. 2b) that misclassifications of tissues fall primarily within developmentally similar microstructures. Misclassifications of villus types are typically confused with other villus types which correspond to similarities in villus growth and branching morphology[3,17]. For example, mature intermediate villi, from which terminal villi grow, are mistaken for terminal villi 53% of the time. Likewise, anchoring villi, which have the same cellular composition as stem villi, are mislabelled as stem villi 21% of the time. Avascular villi, which are commonly associated with the presence of fibrin[59–61], are confounded with fibrin 33% of the time. See Supplementary Fig. 3 for visualisations of predictions across five WSIs of healthy term placentas.

Comparison to perinatal pathologists

To better understand the relative difficulty of reliably identifying different placental tissue microstructures and to further validate our approach, we compare the agreement of four

expert perinatal pathologists in 180 images. Across all tissue microstructure types, pathologists have a moderate agreement score of 0.55 kappa, with low-moderate kappa values for mature intermediate villi (0.468), villus sprouts (0.371), avascular villi (0.154), and anchoring villi (0.051). Pathologists disagree with their majority voted label (Fig. 3a) at least 50% of the time for anchoring villi and avascular villi, highlighting the difficulty of identifying these structures.

Taking the pathologists' majority label as the gold standard, we compare against the ground truth annotations created by C.V. for all 180 images. The resulting kappa value of 0.61 indicates a slightly better agreement than inter-pathologists agreement. For 7/9 tissue microstructures, the pathologists' labels match the ground truth >50% of the time (Fig. 3b). There is a strong match for terminal villi (78%), avascular villi (80%), chorionic plate tissue (90%) and basal plate tissue (99%). Of the two structures with <50% label match, mature intermediate villi (41%) and anchoring villi (27%), these were among the structures with the lowest inter-pathologist agreement as described above.

We contrast model ROC AUC values against pathologist agreement scores for each tissue microstructure type (Fig. 3c). The ROC AUC values have a strong positive correlation (R^2=0.808) with the Cohen's kappa between the pathologists, suggesting the model's predictions are on par with perinatal pathologists for this task. Additionally, pathologists label disagreement (Fig. 3a) show similar patterns to model confusion (Fig. 2b).

Pathologist disagreement in this task is not unexpected as specific, structure-by-structure tissue classification is not part of pathology investigations, partly because this is not humanly feasible at scale. Nonetheless, the 'gestalt' or organised whole, i.e. tissue type and morphology assessment in aggregate is a key part of pathology reporting and disease prediction[33]. Specific tissue classification performance comparable to human experts shows models can accurately quantify placenta biology in a way that is likely relevant to pathology detection. These results highlight the potential of large scale deep learning methods to identify abnormalities in placental microstructures which are too subtle to be recognised by routine light microscopy examination.

Quantitative metrics for placental health

We show how HAPPY can provide new cellular and tissue microstructure quantitative metrics for assessing placental health and we compare these outputs to expectations from our current understanding of placental biology. In Fig. 4a, we present the distribution and variability of predicted cells as a proportion of all cells within a WSI, from five parenchyma WSIs of healthy term placentas. We describe how these predictions reflect the expected internal anatomy of a healthy term placenta. The high proportion of syncytiotrophoblasts (>40%) relative to villus stromal cells matches the expected large surface area to volume ratio of an effective villus tree system optimised for diffusive exchange[17]. The <1% proportion of undifferentiated mesenchymal cells and low proportion of cytotrophoblasts (7%-13%) are characteristic of the late maturation stage of the placenta samples[3,17,33]. The low proportion of leukocytes (2%-4%) are below clinical thresholds for pathological relevance[62].

In Fig. 4b, we present the distribution and variability of predicted tissue microstructures as a proportion of all tissues within a WSI, from five parenchyma WSIs of healthy term placentas. We observe that the proportion of villus microstructures fall within the ranges reported in the literature for healthy term placentas[2,10,17,60,63,64], despite a low number of WSIs evaluated and an expected sampling derived stochasticity. In all but one case, villus structures sit within or a few percent away from expected boundaries, with the anomalous case containing capillaries which had collapsed during slide processing. Excessive regions of fibrin or avascular villi are indicators of pathologic processes[59,65] and the proportions of fibrin and avascular villi are <10% and <2.5%, respectively, which are below the clinical thresholds for pathological relevance[10,59,65,66].

In Fig. 5a, we report the mean cellular compositions for the five chorionic villus types[17] as predicted by the full deep learning pipeline across five WSIs. The predicted cellular proportions match current descriptions and schematics reported in placenta literature[3,17] (Supplementary Table 2). The hierarchical nature of the pipeline reflects the inherent hierarchical relationship between cells and tissues and provides a biologically meaningful way to interpret model predictions. As villus types develop from one another in a tree-like structure (Fig. 5b), their cellular proportions shift along the tree in a continuum and model predictions recapitulate this continuum. For example, the terminal villi, which form the tips of the villus tree and are the primary sites for maternal/fetal diffusive exchange, are characterised by their >50% capillary stromal volume[3,17]. As such, their cellular composition contains the largest proportion of vascular endothelial cells. Conversely, the stem villi, which form the trunk of the villus tree and support the villus structure[3,17], contain a large proportion of structural fibroblasts and vascular myocytes.

A Case Study of Placental Infarction

Placental infarction is a lesion of the placental parenchyma whereby a region or regions of villi undergo ischaemic coagulative necrosis[67]. Caused by a disruption of the maternal circulation within the placental space, it is a key reported finding in pathological investigation as a marker of maternal vascular malperfusion[65,67,68]. At term, when clinically significant[65], it is associated with maternal hypertensive disease, abruption, and fetal growth restriction[69,70].

We compare cell and tissue microstructure predictions between five WSIs of healthy term placentas against the 12 WSIs of term placentas with clinically significant placental infarction. See Supplementary Fig. 4 and Supplementary Fig. 5 for visualisations of cell and tissue microstructure predictions across these slides. Given the biological changes caused by placental infarction[67], at the cellular level we would expect to see fewer syncytiotrophoblast and cytotrophoblast cells and more extravillus trophoblasts and leukocytes. In terms of tissue microstructures, we would expect there to be fewer diffusive villi such as the terminal and mature intermediate villi and, in their place, there should be a larger proportion of fibrin and villi without vasculature (avascular villi). As we do not adjust for the age of the infarction and as younger infarctions will exhibit less nuclear degeneration, we expect these changes to sit along a continuum. We test the significance of these cell and tissue microstructure differences independently between our two groups using Welch's t-test.

Contrasting the proportion of cells across our samples (Fig. 6), we see that syncytiotrophoblast (p=0.0035) and cytotrophoblast cells (p=0.0252) are significantly fewer in placentas with infarction and extravillus trophoblast cells (p=0.0114) and leukocytes (p=0.0016) are significantly higher. In terms of the tissue structures, there are fewer mature intermediate villi (p=0.0111) and more fibrin (p=0.0061) and avascular villi (p=0.0051). Additionally, these proportions of fibrin and avascular villi surpass the healthy expected ranges reported in the literature for 9/12 and 11/12 WSIs with placental infarction, respectively. Similarly, 6/12 and 7/12 WSIs with placental infarction have proportions of terminal villi and mature intermediate villi below expected ranges for healthy term placentas.

Given that placental infarction will result in fewer total nuclei across a slide, we additionally compare the number of predicted cell and tissue microstructure counts per mm^2 area of tissue on the slide (Fig. 7). We estimate this area by splitting the slide into non-overlapping patches and aggregating the area of patches containing at least one nucleus prediction. There are significantly fewer total nuclei (p=0.0031) in the WSIs with placental infarction. We observe additional nominally significant results for both cell and tissue proportions across the slides, however due to limited power in this pilot experiment only syncytiotrophoblast (p=0.0091), leukocyte (p=0.0077) and total nuclei (p=0.0031) pass multiple testing correction.

Discussion

We have presented HAPPY, a three-stage hierarchical deep learning pipeline for quantifying cells and tissue microstructures in healthy human placental histology images. Our method localises and classifies 11 cell types and 9 tissue microstructures across a placenta WSI by following an interpretable biological hierarchy and aggregating cellular community information within tissues. The method's cell and tissue outputs match expectations from perinatal pathologists and placental literature and identifies biologically relevant differences between healthy placentas and those with placental infarction. To our knowledge, this is the first time many of these placental microstructures have been identified across whole slides by automated means. By making tissue predictions at the single cell level, we can quantify microstructures at a sub-patch resolution, without the need for pixel-perfect manual annotations. It is the first exploration of inductive, message-passing graph neural networks to cell graphs for node classification built across entire H&E WSIs. To facilitate further development of graph neural network methodologies suitable for cell graphs, we have released the cell graphs and ground truth tissue data from WSIs of two healthy term placentas for the machine learning research community[71].

Additional improvements could be made to ground truth annotation quality by using immunohistochemical staining of sequential slides. While applications of HAPPY to other organ histology will require ground truth annotations of cells and tissue microstructures specific to that organ, we suggest that the present placenta models could serve as domain relevant pretraining basis for transfer learning. Finally, the HAPPY approach as implemented here has been applied to H&E specifically but could equally be extended to other high content imaging domains.

With HAPPY able to quantify cellular and tissue phenotypes in healthy term placenta WSI samples, we present the foundation for this approach to be developed for extracting quantitative metrics of placental health and pathological processes. Despite having been only trained for inference across healthy samples, HAPPY's outputs meaningfully distinguish between healthy samples and those exhibiting a pathology. Training with WSIs from additional data sources and using tissue microstructures found in earlier gestational ages will improve model generalisability, developmental staging, and provide a more comprehensive view of placenta histology. This can become a valuable digital histopathology tool for perinatal pathologists and placenta research. The placenta is an understudied organ and some of the reasons for this is the labour intensive task of examining placentas, the need for an expert pathologist, and the lack of rich quantitative metrics collected at scale for assessing placenta health.

Already, the HAPPY method can facilitate large-scale morphometric studies of term placenta histology, lending itself to accelerating placenta research and increasing our understanding of the human placenta and its mechanisms.

Methods

At a high level, HAPPY is structured as a supervised deep learning pipeline of three stages: i) nucleus localisation, ii) cell classification and iii) tissue classification (Fig. 1). An image processing module first subsections and rescales the WSI into overlapping patches for the nucleus localisation stage. The nucleus localisation stage uses an object detection model to identify the nuclei within each patch. Using these nuclei coordinates, the image processing module crops patches around each nucleus and inputs them to the cell classification stage. The cell classification stage uses an image classification model to classify the nuclei as belonging to one of 11 placental cell types. These cells are input into the tissue classification stage, becoming nodes within a cell graph across the WSI, the edges of which are built to represent probable cellular interactions within the same structure. This cell graph is input into a node classification graph neural network (GNN) which predicts one of 9 tissue microstructure types to which each cell belongs.

Patient Characteristics, Slide Datasets, and Histological Preparation

The slides used in this work are from a subset of placentas collected as part of routine clinical pathology investigation from two institutes. The first set, from the University of Tartu, Estonia, consists of 110 placentas and 547 slides, collected between 2016-2020 (hereafter referred to as UoT). The use of these pseudo-anonymised samples was approved and the requirement for consent was waived by local ethics committees (approval 289/T-5). The second set, from Hadassah Medical Center, Israel, consists of 200 placentas and 831 slides, collected between 2016-2017 (hereafter referred to as HMC), and the use of these pseudo-anonymised samples was approved and the requirement for consent was waived by local ethics committees (approval 0735-18-HMO).

Of the total samples, we used 7 parenchyma slides from 7 singleton pregnancies of 2nd trimester, preterm and term samples with both healthy placentas and those exhibiting a pathology for training the nucleus localisation and cellular phenotyping stages. For training

the tissue classification stage, we used two parenchyma slides from the placentas of two singleton healthy term pregnancies (Supplementary Table 3). For inference and whole slide cell and tissue microstructure quantification, we used five parenchyma slides from the placentas of five singleton healthy term pregnancies and 12 parenchyma slides from the placentas of eight singleton term pregnancies with placental infarction (Supplementary Table 4). All slides with placental infarction were reported as clinically significant in pathology reports by perinatal pathologists at respective institutes and slides were selected to contain a region of the infarction on the slide.

Given the heterogeneity and resiliency of the placenta, we define 'healthy' as term placentas from pregnancies with no adverse health outcomes during or after pregnancy, a normal birth weight, and from low-risk mothers (non-smokers, no alcohol consumption, non-diabetic, non-hypertensive). Additionally, pathology reports state that histological sections of the parenchyma were 'normal' and that 'villi correspond to gestational age'. The majority of these healthy placentas (4/5) were from pregnancies which had suspected placenta accreta from a 1st-trimester ultrasound, thereby qualifying them for submission to microscopic examination, but with no resulting complications.

Histology slides were prepared using a standard formalin fixing, paraffin-embedded, Hematoxylin and Eosin (H&E) staining procedure. As per clinical guidelines, appropriate, full thickness sites including both chorionic and basal plates were sampled and 5 µm thickness slices were generated. Slides were digitised using a Hamamatsu XR or a 3D HISTECH PANNORAMIC 250 Flash III scanner at x40 magnification.

Image Processing

The image processing module first selects an appropriate slide reading library, one of libvips,[54] openside,[55] or bioformats[56] depending on the slide file format. The slide is partitioned into 1600x1200 pixel (177.44x133.08 µm) patches, with an overlap of 200 (22.18 µm) for the nuclei localisation stage. Patches with mean channel values >245 or <10 are removed to exclude patches containing no tissue. For the cell classification stage, patches of 200x200 pixels (22.18x22.18 µm) are extracted centred on each nucleus. All patches are extracted and loaded onto devices (CPU or GPU) in memory, without the need for additional on-disk storage to store extracted images. Given that slide scanners output different pixel sizes per micrometre, all patches are rescaled to 0.1109 micrometres per pixel. Metadata and results are stored efficiently in an SQLite database and streamed to hdf5 files with the ability to pause and continue partial inference across a WSI.

Nucleus Localisation

The nucleus localisation stage takes 1600x1200 (177.44x133.08 µm) pixel images extracted from the WSI with a 200 (22.18 µm) pixel overlap to ensure all nuclei are shown whole to the model at least once. Locations with duplicate nuclei predictions within a small radius (4 pixels) generated as a result of this overlap are removed with post-processing. We train a RetinaNet[72] with ResNet-101[73] backbone to predict bounding boxes around nuclei in the image, for which centroid coordinates are saved as the final prediction. The model is first fine-tuned from Coco[74] weights for 40 epochs, with an Adam[75] optimiser, focal loss, and a

0.0001 learning rate with a 0.5 decay every 20 epochs. The model with the highest validation F1 score is then fully trained for 60 epochs with a 0.001 learning rate with the same hyperparameters and the model with the highest validation F1 score is saved. Input images are subject to heavy image augmentation, including various H&E-specific stain augmentations (details of augmentation parameters are described in Supplementary Table 5 and Supplementary Fig. 6).

On the validation and test datasets, model performance is evaluated using the F1 score of identified centroids within a certain distance (<3.3 µm) to the ground truth points. This distance is smaller than typical nuclei radii and accounts for minor discrepancies from true centroids in the ground truth data. At inference across a WSI using an NVIDIA A100 GPU, the nucleus localisation stage detects ~1000 nuclei per second.

Cell Classification

The cell classification stage takes 200x200 (22.18x22.18 µm) pixel images with each prior predicted nucleus at its centre and classifies these images into one of 11 placental cell types. These include four trophoblast cells: syncytiotrophoblast, cytotrophoblast, syncytial knot, and extravillus trophoblast; five villus mesenchymal-derived cells: fibroblast, Hofbauer cell, vascular endothelial cell, vascular myocyte, and undifferentiated mesenchymal cell; and two non-villus cells: the maternal decidual cell, and leukocyte. The 22.18 µm radius around each nucleus is large enough to capture the entirety of most cell types in addition to contextual information surrounding the cell which may be relevant for prediction (i.e. red blood cells that are near vascular endothelial cells).

We first fine-tune A ResNet-50[73] model from ImageNet[76] weights for 60 epochs, with an Adam[75] optimiser, cross entropy-loss, and a 0.0001 learning rate with a 0.5 decay every 20 epochs. The model with the highest validation accuracy is fully trained for 100 epochs with the same hyperparameters and the model with the highest validation accuracy is saved. Due to class imbalance, minority classes are oversampled during training to balance the distribution. Images input to the classification model are subject to the same image augmentation as the images used for nuclei detection (Supplementary Table 5). As a final post-processing step, a k-d tree is constructed across all syncytial knot predictions to convert isolated knots into syncytiotrophoblasts and to group clusters of syncytial knot nuclei into a single point. Syncytial knot nuclei which have fewer than 4 neighbours within a 50-pixel radius are relabelled as syncytiotrophoblasts and groups with 4 or more neighbours[77] have their neighbours removed. At inference across a WSI using an NVIDIA A100 GPU, the cell classification stage classifies ~230 cells per second.

Tissue Classification

The tissue classification stage comprises cell graph construction and supervised graph neural network node classification for each cell into one of nine placental tissue types. These include the term chorionic villus types: stem villi, anchoring villi, mature intermediate villi, terminal villi, and villus sprouts, the maternal/fetal surfaces: the chorionic plate and basal plate/septum, and areas which in large quantities would indicate pathology[10,59,60,66]: fibrin and avascular villi.

For the cell graph, we define nodes from the outputs of the cellular phenotyping stage, with node features comprising the 64-dimension embedding vectors from the penultimate layer of the cell classifier. The undirected edges connecting the cell nodes are constructed from the intersection of two other edge-building algorithms, k-nearest neighbours (k=5)[78] and Delaunay Triangulation[79], with the addition of self-loops. This intersection benefits from the more sparsely connected Delaunay Triangulation graph but limits the number of edges which could cross from one tissue boundary to another. This graph construction allows a message-passing model to aggregate cellular information within distinct tissue microstructures while accounting for differences in tissue size and internal cell distances.

We train a randomly initialised, inductive, ClusterGCN[80] model with 16 GraphSAGEConv[81] layers each with 256 hidden units for 2000 epochs with an Adam[75] optimiser, custom weighted cross entropy-loss, a 0.001 learning rate, a batch size of 200 with batch normalisation, and a subgraph sampling size of 400 neighbours. The model with the highest validation accuracy, calculated without neighbourhood sampling and for an intersection graph with k=8, is saved as the final model. For each node the message-passing algorithm samples and aggregates the node features of nearby nodes up to 16 edge connections away. These aggregated node features are used by the model to predict the tissue type of that node. In this way, the model is aggregating the cellular community which comprises each tissue microstructure to make its prediction. At inference across a WSI using a laptop CPU, the tissue classification stage classifies ~4,500 nodes per second.

A key benefit of the node classification approach is the freedom for the model to assign different tissue types to different sections of the same continuous structure. Placental tissues grow from one another in a tree-like morphology so any cross-sectional cut of what appears to be a single structure may contain multiple valid classifications. For example, a cut of a mature intermediate villus is likely to have terminal villi growths but may also contain fibrin, resulting in three different classes. Additionally, the distinctions between villus types are not necessarily discrete; a terminal villus is distinguished from a mature intermediate villus by having >50% of its stroma taken up by capillaries and by having vasculosyncitial membranes[3,17,33]. However, a section of a mature intermediate villus which has unusually many capillaries but no vasculosyncitial membranes, for example, might be (mis)classified by the model as a terminal villus section but is not necessarily incorrect.

Datasets and Data Annotation

Ground truth training, validation and test dataset annotation were performed using QuPath[82]. To efficiently train models with a relatively small number of data points (<17k), datasets for the nuclei localisation and cell classification stages were created iteratively by bootstrapping and correcting prior models' predictions in patches of new, unseen slides. Training, validation, and test splits (see Supplementary Table 6) were randomly generated at around a 70/15/15% ratio for each dataset and model performance was evaluated both individually for datasets and in combination.

For the tissue classification stage, ground truth annotations were created by drawing rough boundaries around tissue microstructures, the class of which was assigned to nuclei nodes

within the boundary. Validation and test regions of the slide were explicitly chosen such that they contained a similar distribution of tissue microstructures to the training set (See Supplementary Fig. 7). Splitting datasets by region was found to limit information leakage when different nuclei of same tissues were shared across dataset splits. This is in contrast to many graph learning datasets which generate dataset splits randomly across nodes.

Comparison to Perinatal Pathologists

To assess the accuracy of ground truth annotations and to judge the relative difficulty of identifying placental tissue types, four practising expert perinatal pathologists performed a similar labelling task across tissue microstructures. Each pathologist was shown a series of images containing a centred tissue microstructure out of 180 total images with the task of labelling the tissue type of that centred structure. Images were generated from a random, class-balanced subset of the ground truth annotations and pathologists were blind to the original annotation and each other's labels. Images were cropped to display some contextual background, similar to the context a 16-layer message-passing GNN may see, however, it was expected that tissue types were identifiable from their cellular composition alone.

Participants were first presented with a Standard Operating Procedure (SOP) and a tutorial (Supplementary Files 1 and 2). These documents detailed the labelling setup, what data would be collected, how it would be used, and included links to current literature relevant to placental tissue microstructures. Participants were invited to a project in the browser-based software LabelBox[83] where they were sequentially presented with images and could choose one of 12 tissue types for that image. Alternatively, they could state that the type was *Unclear* or *Not listed* and they could leave a comment. Participants were informed that all images came from a healthy term placenta. After completion, their tissue type labels were compared for Cohen's kappa[84] agreement scores against each other, the original annotations, and the model.

Hardware, Software and Libraries

All training and inference were performed on a single NVIDIA A100 GPU. Whole slide image processing readers were either libvips (v8.9.2)[54], openslide (v3.4.1)[55], or bioformats (v4.0.4)[56] with libvips taking priority. Code was written in Python (v3.7.2) with the PyTorch (v1.1.1)[85] deep learning framework for the nucleus localisation and cellular phenotyping pipelines and the PyTorch Geometric extension (v2.0.4)[86] for the tissue phenotyping pipeline. Results were recorded in an SQLite database (v3.2.7) and hdf5 files (v3.2.1). WSI visualisation and cell and tissue annotations were done in QuPath (v0.3.1)[82].

Data availability

The datasets generated and analysed for training each deep learning model are available at: https://tinyurl.com/happyplacenta. Restrictions apply to the availability of the placenta histology in-house data, which were used with rigorous ethical and legal approval by local ethics boards for the current study, and are thus not publicly available. All requests for academic use of analysed data can be addressed to the corresponding authors. Source data for Figs. 2, 3, 4, 5, 6, 7 are provided with the paper.

Code availability

Code is available at the following GitHub repository: https://github.com/Nellaker-group/happy

Acknowledgements

Claudia Vanea is supported by the EPSRC Center for Doctoral Training in Health Data Science (EP/S02428X/1). Cecilia M. Lindgren is supported by the Li Ka Shing Foundation, NIHR Oxford Biomedical Research Centre, Oxford, NIH (1P50HD104224-01), Gates Foundation (INV-024200), and a Wellcome Trust Investigator Award (221782/Z/20/Z). Triin Laisk was funded by the European Regional Development Fund and the programme Mobilitas Pluss (MOBTP155) and the Estonian Research Council grant PSG776.

The computational aspects of this research were supported by the Wellcome Trust Core Award Grant Number 203141/Z/16/Z and the NIHR Oxford BRC. The views expressed are those of the author(s) and not necessarily those of the NHS, the NIHR or the Department of Health.

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS. The data used for supplementary analyses described in this manuscript were obtained from: the GTEx Portal on 24/10/22.

Contributions

C.V., C.M.L. and C.N. conceived the study and designed the experiments. C.V. and C.N. wrote the code, C.V. created ground truth annotations and performed the experimental analysis. J.D., V.R., O.D., L.S., K.M., D.H., H.H., and T.L. curated the in-house datasets. S.S., K.M., W.T.P., L.M.E., partook in the pathologist comparison. C.V., C.M.L. and C.N. prepared the manuscript with input, revisions, and approval from J.D., V.R., O.D., S.S., L.S., K.M., W.T.P., D.H., A.F., H.H., T.L., and L.M.E.. L.M.E. mentored C.V. on placental biology and perinatal pathology. C.N. and C.M.L. supervised the research.

Figures

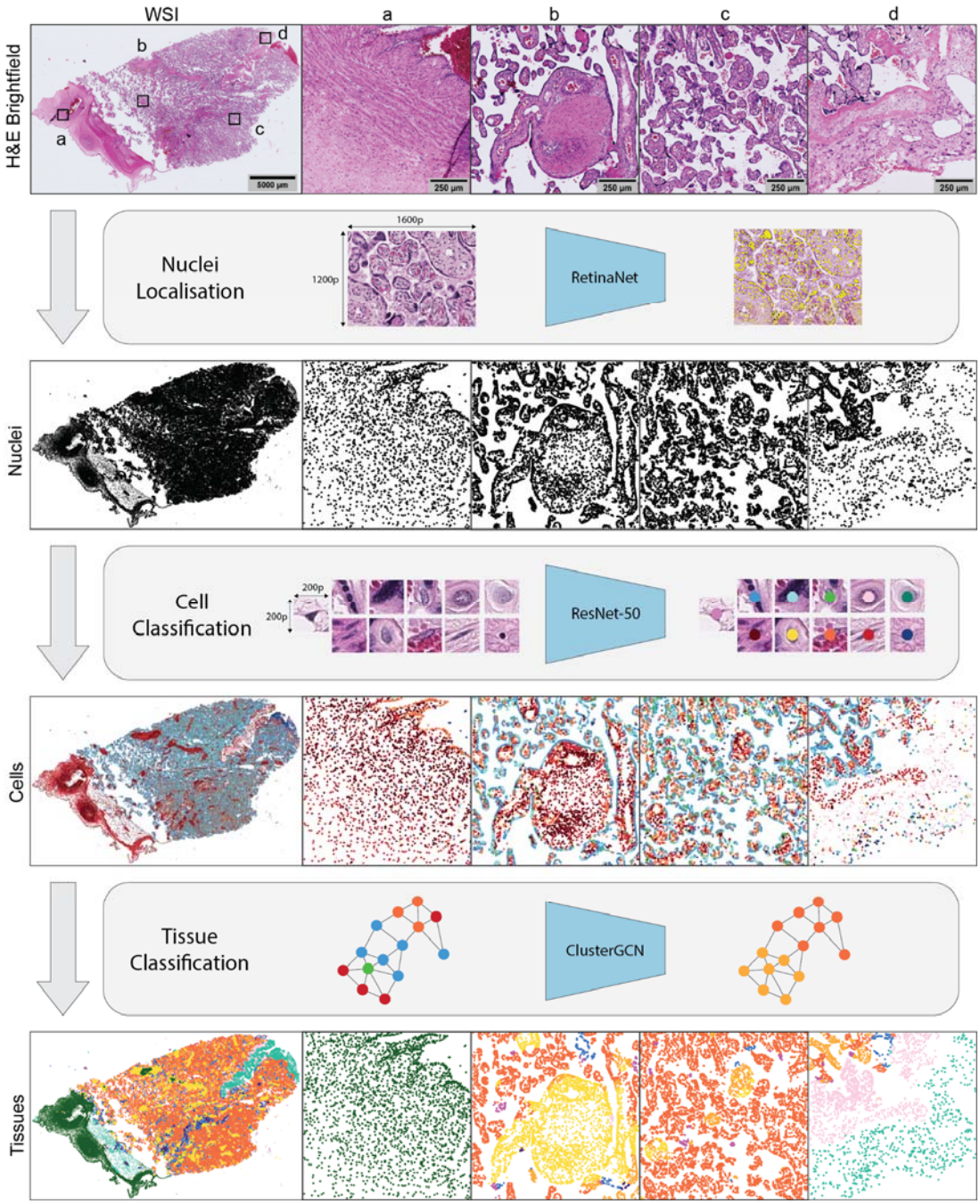

Fig. 1: HAPPY workflow. A whole slide image (WSI) is first sectioned into overlapping 1600x1200 (177.44x133.08 µm) pixel images and passed to an object detection RetinaNet model which identifies the nuclei in these images. 200x200 (22.18x22.18 µm) pixel images centred on each nucleus are classified into one of 11 cell types by a ResNet-50 model. The 64-dimension embeddings from the cell classifier and their corresponding nucleus coordinates are used to build a cell graph across the whole slide image. The cell graph is input into a ClusterGCN graph neural network which classifies the tissue microstructure to which each cell belongs. Images a-d show characteristic tissue regions of the WSI: (a) chorionic plate, (b) stem and distal villi, (c) distal villi, (d) basal plate and anchoring villi.

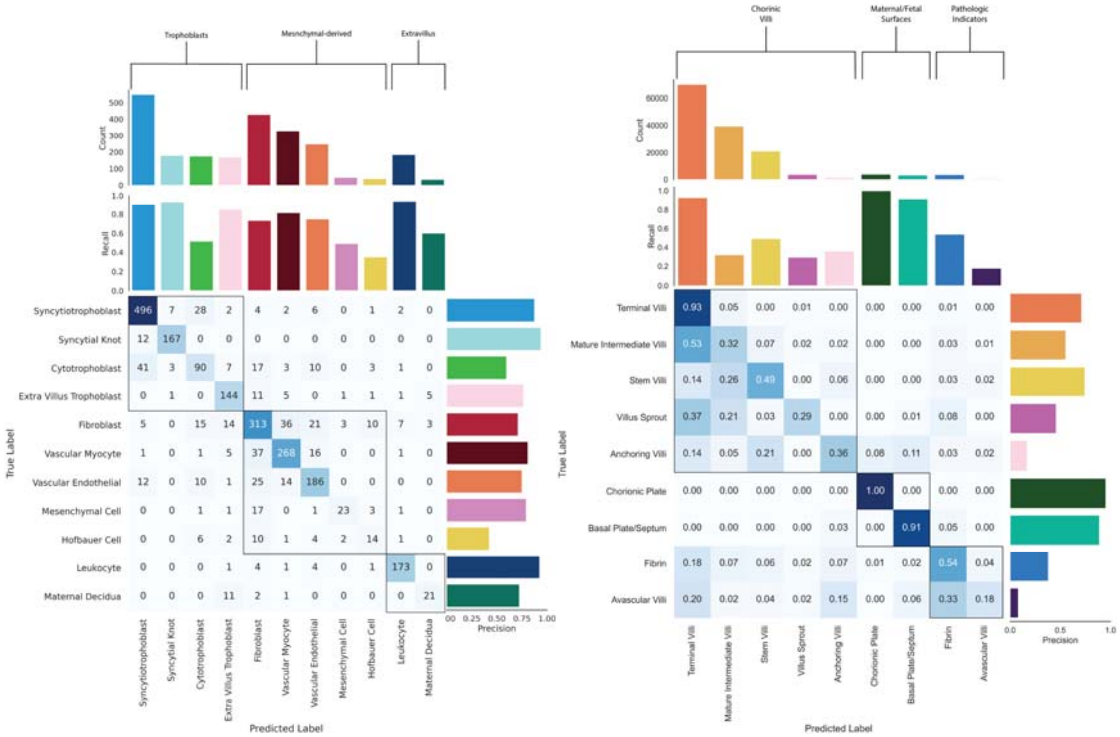

Fig. 2: Cell and tissue classifier performance on unseen test data. (a) Confusion matrix and precision and recall values of cell classifier predictions. Cell types are clustered into categories (trophoblast, mesenchymal-derived, extravillus) and ordered by their counts. Categories are highlighted by the topmost brackets and squares within the confusion matrix. (b) Confusion matrix and precision and recall values of tissue classifier predictions. Confusion matrix values are represented as a proportion of predictions relative to the number of samples per tissue type. Tissue types are clustered into categories (chorionic villi, maternal/fetal surfaces, pathologic indicators) and ordered by their counts. Categories are highlighted by the topmost brackets and squares within the confusion matrix.

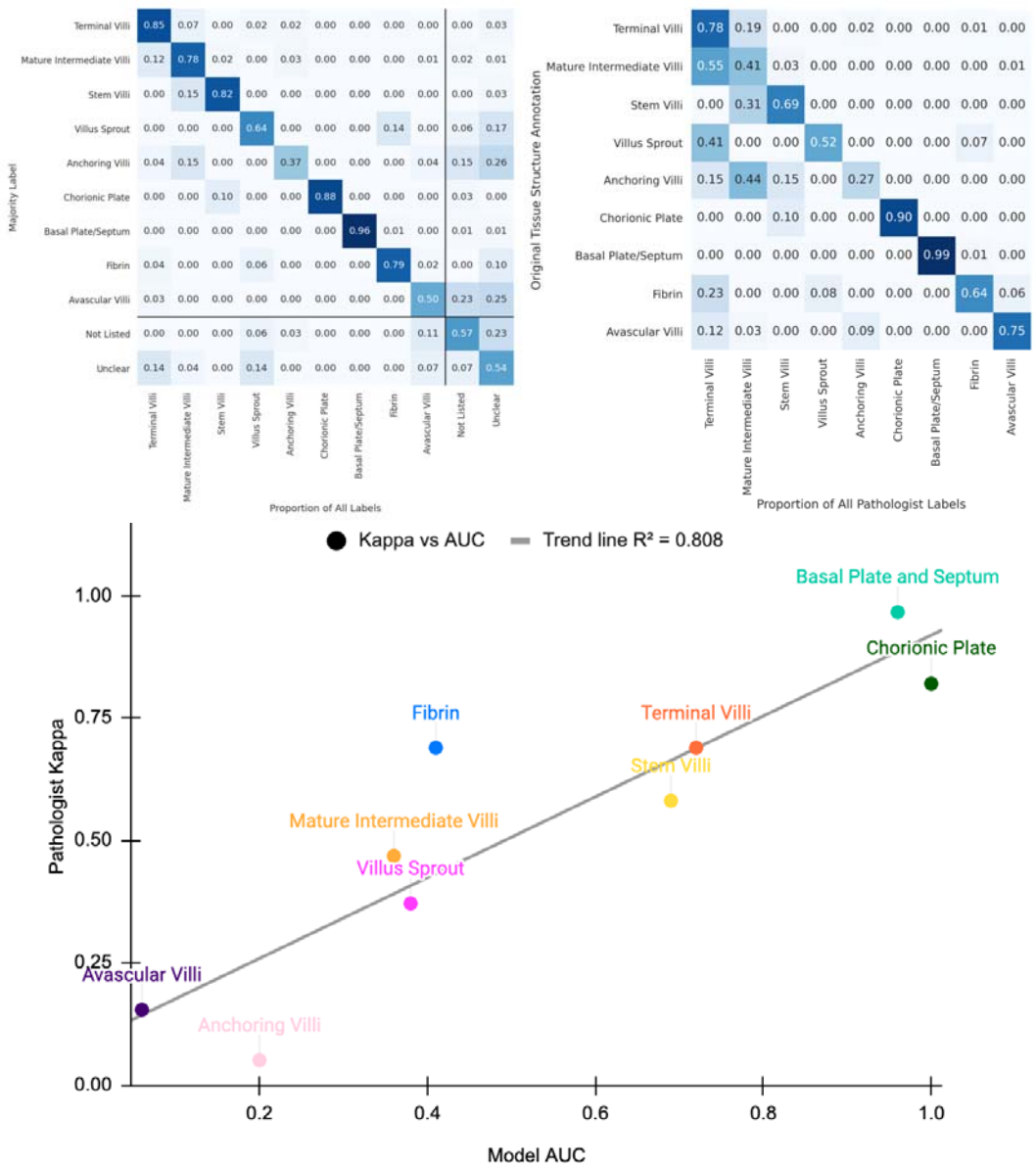

Fig. 3: Agreement and model confusion across tissue types. (a) Inter-pathologist confusion matrix showing for each image with a majority pathologist label (y-axis) how much variance there was against that majority (x-axis). Lines indicate the additional *Not Listed* and *Unclear* choices offered to pathologists. (b) Confusion matrix which shows the proportion of matching pathologists' labels with ground truth annotations. (c) Model performance ROC AUC scores plotted against pathologist Kappa values by tissue type showing a strong positive correlation (R^2 = 0.808).

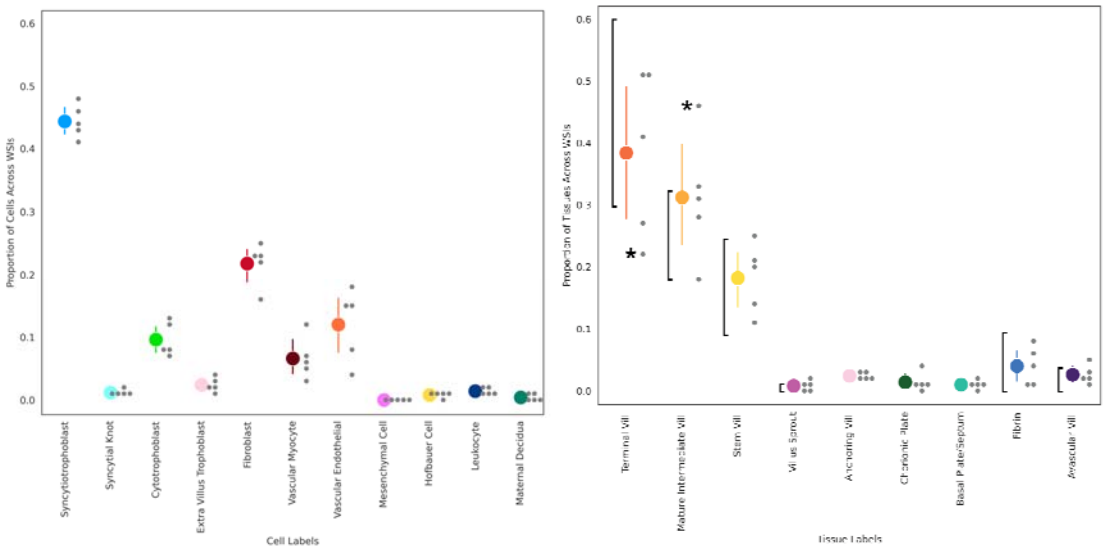

Fig. 4: Predicted cell and tissue type proportions across five WSIs from healthy term placentas. (a) Cell proportions across WSIs. (b) Tissue proportions across WSIs with vertical braces showing the expected ranges reported in the literature for healthy term placentas[10,17,33,60,63–65]. Coloured dots represent mean proportions with standard error bars and grey dots show individual data points. Asterisk highlights the anomalous terminal villi and mature intermediate villi proportions for one WSI which had collapsed capillaries.

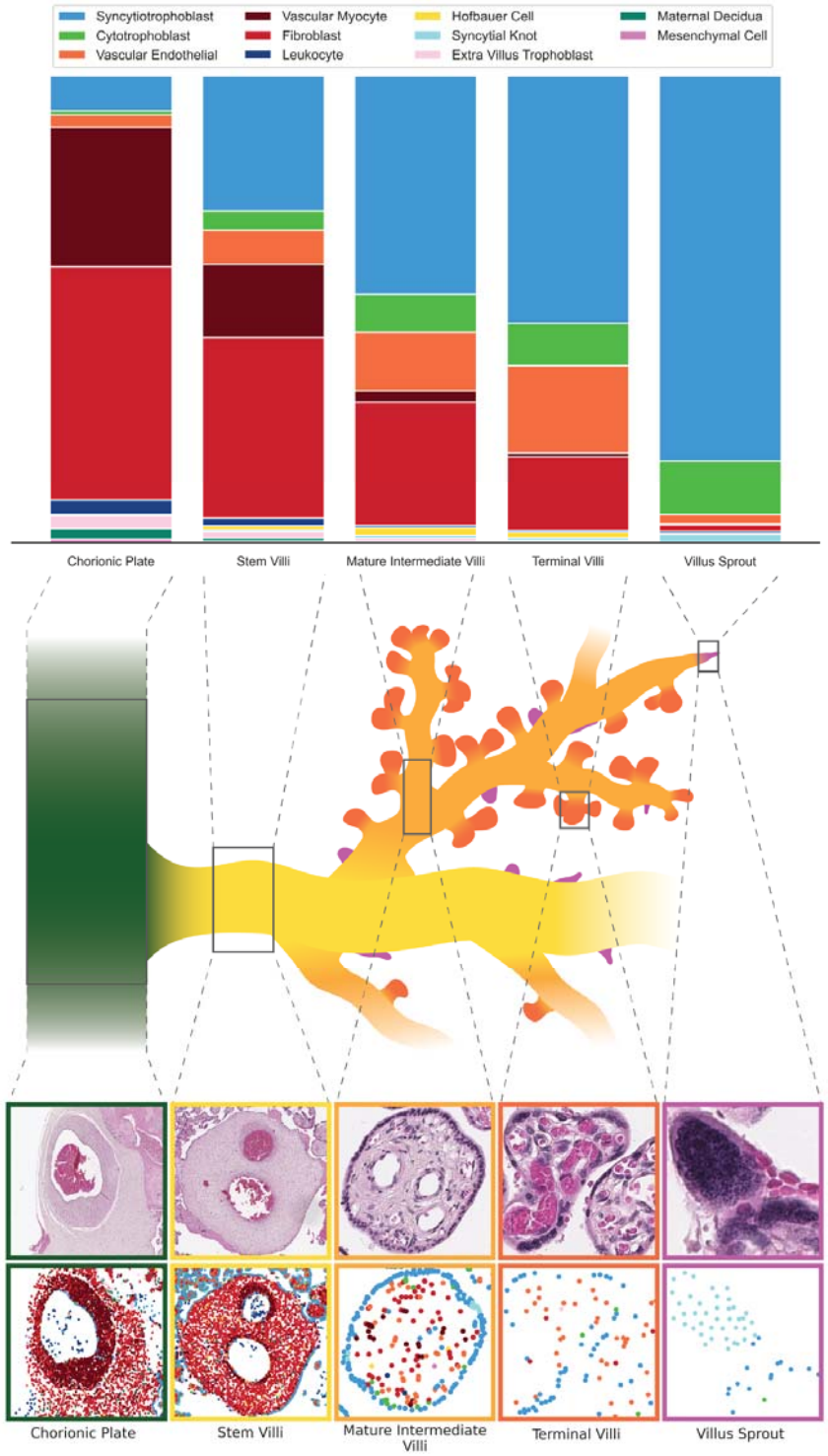

Fig. 5: Proportions of predicted cell types within predicted term chorionic villus tissue microstructures and their corresponding locations in a villus tree schematic. (a) The mean proportion of predicted cell types within each predicted term chorionic villus tissue microstructure for five WSI. (b) A villus tree schematic showing how each villus structure relates to and grows from other villus structures. (c) Examples from histology for each villus structure with the top row containing the raw histology and the bottom row containing cell predictions.

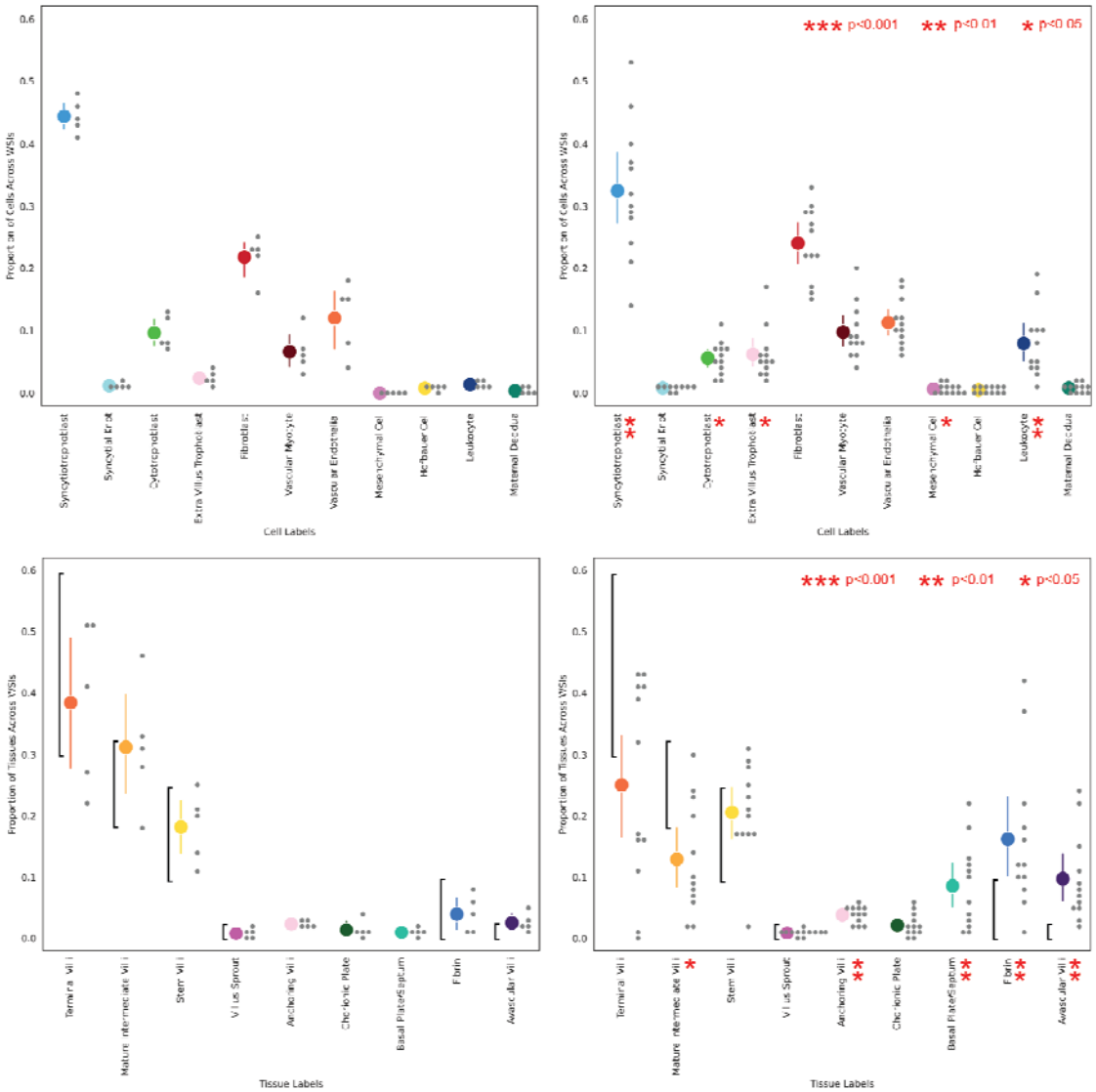

Fig. 6: Predicted cell and tissue microstructure proportions across five WSIs of healthy term placentas and 12 WSIs of placentas with placental infarction. (a) Cell proportions across healthy term placentas, (b) cell proportions across term placentas with placental infarction, (c) tissue proportions across healthy term placentas, (d) tissue proportions across term placentas with placental infarction. Coloured dots represent mean proportions with standard error bars and grey dots show individual data points. Expected healthy ranges for tissue microstructures, as reported in the literature, are shown by black vertical bars. Nominal significant differences in cell and tissue structures between the two groups are calculated using Welch's t-test and shown by red asterisks.

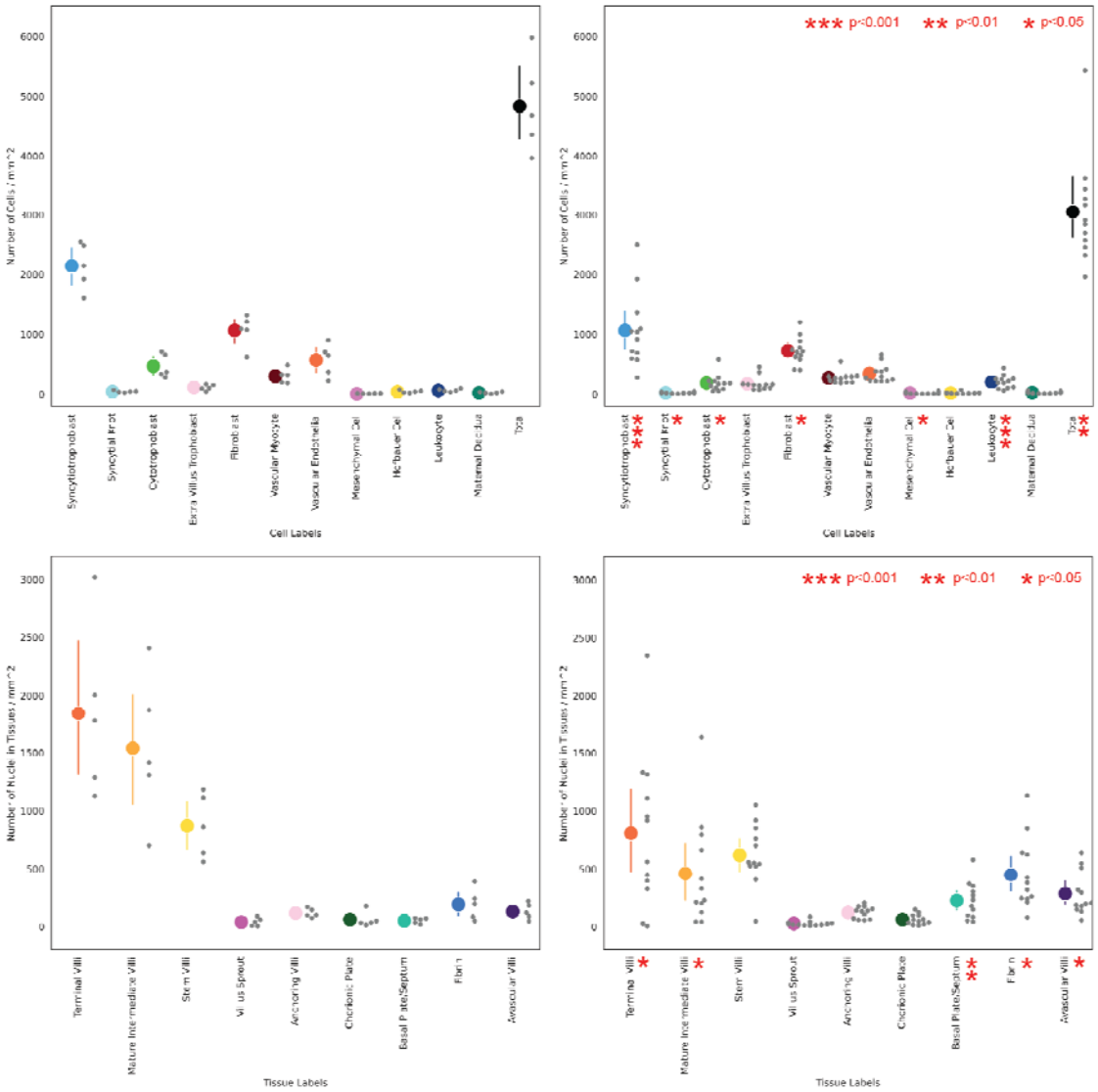

Fig. 7: Predicted cell and tissue microstructure density per area mm^2 across five WSIs of healthy term placentas and 12 WSIs of placentas with placental infarction. (a) Cell density across healthy term placentas including total nuclei, (b) cell density across term placentas with placental infarction including total nuclei, (c) tissue node density across healthy term placentas, (d) tissue node density across term placentas with placental infarction. Coloured dots represent mean proportions with standard error bars and grey dots show individual data points. Nominal significant differences in cell and tissue structures between the two groups are calculated using Welch's t-test and shown by red asterisks.

Tables

Table 1: Summary of performance on unseen test data for each deep learning stage.

	F1		
Nuclei Detection	0.884		
	Accuracy	**Top-2 Accuracy**	**ROC AUC**
Cell Classification	80%	93%	0.969
Tissue Classification	67%	90%	0.898

References

1. Morgan, T. K. Frontiers in Placental Pathology. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 379–382 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_58.
2. Turowski, G., Arbuckle, S. & Parks, W. T. The Placental Pathology Report. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 371–377 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_57.
3. Ernst, L. M. & Carreon, C. K. Placenta. in *Color Atlas of Human Fetal and Neonatal Histology* (eds. Ernst, L. M., Ruchelli, E. D., Carreon, C. K. & Huff, D. S.) 399–424 (Springer International Publishing, 2019). doi:10.1007/978-3-030-11425-1_36.
4. Burton, G. J., Fowden, A. L. & Thornburg, K. L. Placental Origins of Chronic Disease. *Physiol. Rev.* **96**, 1509–1565 (2016).
5. Gordijn, S. J., Dahlstrom, J. E., Khong, T. Y. & Ellwood, D. A. Histopathological examination of the placenta: key issues for pathologists and obstetricians. *Pathology (Phila.)* **40**, 176–179 (2008).
6. Reyes, L. & Golos, T. G. Hofbauer Cells: Their Role in Healthy and Complicated Pregnancy. *Front. Immunol.* **9**, 2628 (2018).
7. Taylor, L. A., Gallagher, K., Ott, K. A. & Gernand, A. D. How often is the placenta included in human pregnancy research? A rapid systematic review of the literature. Preprint at https://doi.org/10.12688/gatesopenres.13215.1 (2021).
8. Roberts, D. J. Perinatal Pathology: Practice Suggestions for Limited–Resource Settings. *Arch. Pathol. Lab. Med.* **137**, 775–781 (2013).
9. Hargitai, B., Marton, T. & Heerema-McKenney, A. Indications for Examining the Placenta. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 31–37 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_3.

10. Hargitai, B., Marton, T. & Cox, P. M. BEST PRACTICE NO 178□: Examination of the human placenta. *J. Clin. Pathol.* **57**, 785–792 (2004).

11. Papadogiannakis, N. How can we ensure that placental pathology plays a greater role in treating neonatal morbidity? *Acta Paediatr.* **109**, 1514–1515 (2020).

12. Turowski, G., Tony Parks, W., Arbuckle, S., Jacobsen, A. F. & Heazell, A. The structure and utility of the placental pathology report. *APMIS* **126**, 638–646 (2018).

13. Goldstein, J. A., Gallagher, K., Beck, C., Kumar, R. & Gernand, A. D. Maternal-Fetal Inflammation in the Placenta and the Developmental Origins of Health and Disease. *Front. Immunol.* **11**, (2020).

14. Barker, D. J. P. & Thornburg, K. L. Placental programming of chronic diseases, cancer and lifespan: a review. *Placenta* **34**, 841–845 (2013).

15. Heerema-McKenney, A. Placental Development with Expected Normal Gross and Microscopic Findings. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 9–29 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_2.

16. Mobadersany, P., Cooper, L. A. D. & Goldstein, J. A. GestAltNet: aggregation and attention to improve deep learning of gestational age from placental whole-slide images. *Lab. Invest.* **101**, 942–951 (2021).

17. Benirschke, K., Burton, G. J. & Baergen, R. N. Architecture of Normal Villous Trees. in *Pathology of the Human Placenta* (eds. Benirschke, K., Burton, G. J. & Baergen, R. N.) 101–144 (Springer, 2012). doi:10.1007/978-3-642-23941-0_7.

18. Khong, T. Y. *et al.* Observer reliability in assessing placental maturity by histology. *J. Clin. Pathol.* **48**, 420–423 (1995).

19. Al-Adnani, M., Marnerides, A., George, S., Nasir, A. & Weber, M. A. “Delayed Villous Maturation” in Placental Reporting: Concordance among Consultant Pediatric Pathologists at a Single Specialist Center. *Pediatr. Dev. Pathol.* **18**, 375–379 (2015).

20. Redline, R. W. *et al.* Maternal vascular underperfusion: nosology and reproducibility of placental reaction patterns. *Pediatr. Dev. Pathol. Off. J. Soc. Pediatr. Pathol. Paediatr.*

Pathol. Soc. **7**, 237–249 (2004).

21. Sun, C.-C. J., Revell, V. O., Belli, A. J. & Viscardi, R. M. Discrepancy in pathologic diagnosis of placental lesions. *Arch. Pathol. Lab. Med.* **126**, 706–709 (2002).

22. Custer, J. W. *et al.* Diagnostic errors in the pediatric and neonatal ICU: a systematic review. *Pediatr. Crit. Care Med. J. Soc. Crit. Care Med. World Fed. Pediatr. Intensive Crit. Care Soc.* **16**, 29–36 (2015).

23. Li, R., Yao, J., Zhu, X., Li, Y. & Huang, J. Graph CNN for Survival Analysis on Whole Slide Pathological Images. in *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018* (eds. Frangi, A. F., Schnabel, J. A., Davatzikos, C., Alberola-López, C. & Fichtinger, G.) 174–182 (Springer International Publishing, 2018). doi:10.1007/978-3-030-00934-2_20.

24. Chen, R. J. *et al.* Pathomic Fusion: An Integrated Framework for Fusing Histopathology and Genomic Features for Cancer Diagnosis and Prognosis. *IEEE Trans. Med. Imaging* **PP**, (2020).

25. Chen, R. J. *et al.* Scaling Vision Transformers to Gigapixel Images via Hierarchical Self-Supervised Learning. in *2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* 16123–16134 (IEEE, 2022). doi:10.1109/CVPR52688.2022.01567.

26. Lu, M. Y. *et al.* Data Efficient and Weakly Supervised Computational Pathology on Whole Slide Images. *ArXiv200409666 Cs Eess Q-Bio* (2020).

27. Sirinukunwattana, K. *et al.* Novel digital signatures of tissue phenotypes for predicting distant metastasis in colorectal cancer. *Sci. Rep.* **8**, 13692 (2018).

28. Keren, L. *et al.* A Structured Tumor-Immune Microenvironment in Triple Negative Breast Cancer Revealed by Multiplexed Ion Beam Imaging. *Cell* **174**, 1373-1387.e19 (2018).

29. Lu, W. *et al.* SlideGraph+: Whole slide image level graphs to predict HER2 status in breast cancer. *Med. Image Anal.* **80**, 102486 (2022).

30. Bejnordi, B. E. *et al.* Using deep convolutional neural networks to identify and classify

tumor-associated stroma in diagnostic breast biopsies. *Mod. Pathol. Off. J. U. S. Can. Acad. Pathol. Inc* **31**, 1502–1512 (2018).

31. Ferlaino, M. *et al.* Towards Deep Cellular Phenotyping in Placental Histology. *ArXiv180403270 Cs* (2018).

32. Clymer, D. *et al.* Decidual Vasculopathy Identification in Whole Slide Images Using Multiresolution Hierarchical Convolutional Neural Networks. *Am. J. Pathol.* **190**, 2111–2122 (2020).

33. Turowski, G., Mooney, E. E. & Scheimberg, I. B. Maturity Anomalies. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 115–121 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_14.

34. Jahanifar, M., Tajeddin, N. Z., Koohbanani, N. A. & Rajpoot, N. Robust Interactive Semantic Segmentation of Pathology Images with Minimal User Input. in *2021 IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)* 674–683 (IEEE, 2021). doi:10.1109/ICCVW54120.2021.00081.

35. Graham, S. *et al.* MILD-Net: Minimal information loss dilated network for gland instance segmentation in colon histology images. *Med. Image Anal.* **52**, 199–211 (2019).

36. Khened, M., Kori, A., Rajkumar, H., Krishnamurthi, G. & Srinivasan, B. A generalized deep learning framework for whole-slide image segmentation and analysis. *Sci. Rep.* **11**, 11579 (2021).

37. Sirinukunwattana, K. *et al.* Gland segmentation in colon histology images: The glas challenge contest. *Med. Image Anal.* **35**, 489–502 (2017).

38. Kiemen, A. L. *et al.* CODA: quantitative 3D reconstruction of large tissues at cellular resolution. *Nat. Methods* **19**, 1490–1499 (2022).

39. Xu, Y. *et al.* Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features. *BMC Bioinformatics* **18**, 281 (2017).

40. Hou, L. *et al.* Patch-Based Convolutional Neural Network for Whole Slide Tissue

Image Classification. in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2424–2433 (2016). doi:10.1109/CVPR.2016.266.

41. Xu, Y., Zhu, J.-Y., Chang, E. I.-C., Lai, M. & Tu, Z. Weakly supervised histopathology cancer image segmentation and classification. *Med. Image Anal.* **18**, 591–604 (2014).

42. Hou, L. *et al.* Patch-based Convolutional Neural Network for Whole Slide Tissue Image Classification. Preprint at https://doi.org/10.48550/arXiv.1504.07947 (2016).

43. Li, M. M., Huang, K. & Zitnik, M. Graph representation learning in biomedicine and healthcare. *Nat. Biomed. Eng.* 1–17 (2022) doi:10.1038/s41551-022-00942-x.

44. Lu, W., Graham, S., Bilal, M., Rajpoot, N. & Minhas, F. Capturing Cellular Topology in Multi-Gigapixel Pathology Images. in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)* 1049–1058 (IEEE, 2020). doi:10.1109/CVPRW50498.2020.00138.

45. Li, J., Chen, S., Pan, X., Yuan, Y. & Shen, H.-B. Cell clustering for spatial transcriptomics data with graph neural networks. *Nat. Comput. Sci.* **2**, 399–408 (2022).

46. Javed, S. *et al.* Cellular community detection for tissue phenotyping in colorectal cancer histology images. *Med. Image Anal.* **63**, 101696 (2020).

47. Zhou, Y. *et al.* CGC-Net: Cell Graph Convolutional Network for Grading of Colorectal Cancer Histology Images. *ArXiv190901068 Cs Eess* (2019).

48. Gadiya, S., Anand, D. & Sethi, A. Histographs: Graphs in Histopathology. *ArXiv190805020 Cs Eess* (2019).

49. Stassen, S. V. *et al.* PARC: ultrafast and accurate clustering of phenotypic data of millions of single cells. *Bioinformatics* **36**, 2778–2786 (2020).

50. Jaume, G. *et al.* Towards Explainable Graph Representations in Digital Pathology. Preprint at https://doi.org/10.48550/arXiv.2007.00311 (2020).

51. Kim, J. *et al.* Unsupervised discovery of tissue architecture in multiplexed imaging. *Nat. Methods* 1–9 (2022) doi:10.1038/s41592-022-01657-2.

52. Wang, J., Chen, R. J., Lu, M. Y., Baras, A. & Mahmood, F. Weakly Supervised Prostate Tma Classification Via Graph Convolutional Networks. in *ISBI 2020 - 2020 IEEE*

International Symposium on Biomedical Imaging 239–243 (IEEE Computer Society, 2020). doi:10.1109/ISBI45749.2020.9098534.

53. Brbić, M. *et al.* Annotation of spatially resolved single-cell data with STELLAR. *Nat. Methods* 1–8 (2022) doi:10.1038/s41592-022-01651-8.

54. Martinez, K. & Cupitt, J. VIPS - a highly tuned image processing software architecture. in *IEEE International Conference on Image Processing 2005* vol. 2 II–574 (2005).

55. Goode, A., Gilbert, B., Harkes, J., Jukic, D. & Satyanarayanan, M. OpenSlide: A vendor-neutral software foundation for digital pathology. *J. Pathol. Inform.* **4**, 27 (2013).

56. Linkert, M. *et al.* Metadata matters: access to image data in the real world. *J. Cell Biol.* **189**, 777–782 (2010).

57. Graham, S. *et al.* HoVer-Net: Simultaneous Segmentation and Classification of Nuclei in Multi-Tissue Histology Images. Preprint at http://arxiv.org/abs/1812.06499 (2019).

58. Gamper, J. *et al.* PanNuke Dataset Extension, Insights and Baselines. Preprint at http://arxiv.org/abs/2003.10778 (2020).

59. Katzman, P. J., Ernst, L. M. & Scheimberg, I. B. Massive Perivillous Fibrinoid Deposition and Maternal Floor Infarct. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 77–82 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_8.

60. Redline, R. W. & Pappin, A. Fetal thrombotic vasculopathy: The clinical significance of extensive avascular villi. *Hum. Pathol.* **26**, 80–85 (1995).

61. Redline, R. W. & Ravishankar, S. Fetal vascular malperfusion, an update. *APMIS* **126**, 561–569 (2018).

62. Katzman, P. J. Chorionic Plate Chronic Inflammatory Lesions Including Eosinophilic/T-Cell Chorionic Vasculitis. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 109–113 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_13.

63. Castellucci, M., Scheper, M., Scheffen, I., Celona, A. & Kaufmann, P. The development of the human placental villous tree. *Anat. Embryol. (Berl.)* **181**, 117–128 (1990).

64. Demir, R., Kosanke, G., Kohnen, G., Kertschanska, S. & Kaufmann, P. Classification of human placental stem villi: Review of structural and functional aspects. *Microsc. Res. Tech.* **38**, 29–41 (1997).

65. Khong, T. Y. *et al.* Sampling and Definitions of Placental Lesions: Amsterdam Placental Workshop Group Consensus Statement. *Arch. Pathol. Lab. Med.* **140**, 698–713 (2016).

66. Rogers, B. B. *et al.* Avascular Villi, Increased Syncytial Knots, and Hypervascular Villi Are Associated with Pregnancies Complicated by Factor V Leiden Mutation. *Pediatr. Dev. Pathol. Off. J. Soc. Pediatr. Pathol. Paediatr. Pathol. Soc.* **13**, 341–347 (2010).

67. Fitzgerald, B. Infarction. in *Pathology of the Placenta: A Practical Guide* (eds. Khong, T. Y., Mooney, E. E., Nikkels, P. G. J., Morgan, T. K. & Gordijn, S. J.) 57–65 (Springer International Publishing, 2019). doi:10.1007/978-3-319-97214-5_6.

68. Freedman, A. A., Keenan-Devlin, L. S., Borders, A., Miller, G. E. & Ernst, L. M. Formulating a Meaningful and Comprehensive Placental Phenotypic Classification. *Pediatr. Dev. Pathol. Off. J. Soc. Pediatr. Pathol. Paediatr. Pathol. Soc.* **24**, 337–350 (2021).

69. Mousa, H. A. & Alfirevic1, Z. Do placental lesions reflect thrombophilia state in women with adverse pregnancy outcome? *Hum. Reprod.* **15**, 1830–1833 (2000).

70. Becroft, D. M. O., Thompson, J. M. D. & Mitchell, E. A. The Epidemiology of Placental Infarction at Term. *Placenta* **23**, 343–351 (2002).

71. Vanea, C. *et al.* A New Graph Node Classification Benchmark: Learning Structure from Histology Cell Graphs. in (New Frontiers in Graph Learning at NeurIPS, 2022). doi:10.48550/arXiv.2211.06292.

72. Lin, T.-Y., Goyal, P., Girshick, R., He, K. & Dollár, P. Focal Loss for Dense Object Detection. *ArXiv170802002 Cs* (2018).

73. He, K., Zhang, X., Ren, S. & Sun, J. Deep Residual Learning for Image Recognition. *ArXiv151203385 Cs* (2015).

74. Lin, T.-Y. *et al.* Microsoft COCO: Common Objects in Context. *ArXiv14050312 Cs* (2015).

75. Kingma, D. P. & Ba, J. Adam: A Method for Stochastic Optimization. in (2014).

76. Deng, J. *et al.* ImageNet: A large-scale hierarchical image database. in *2009 IEEE Conference on Computer Vision and Pattern Recognition* 248–255 (2009). doi:10.1109/CVPR.2009.5206848.

77. Loukeris, K., Sela, R. & Baergen, R. N. Syncytial knots as a reflection of placental maturity: reference values for 20 to 40 weeks' gestational age. *Pediatr. Dev. Pathol. Off. J. Soc. Pediatr. Pathol. Paediatr. Pathol. Soc.* **13**, 305–309 (2010).

78. Eppstein, D., Paterson, M. S. & Yao, F. F. On Nearest-Neighbor Graphs. *Discrete Comput. Geom.* **17**, 263–282 (1997).

79. Guibas, L. J., Knuth, D. E. & Sharir, M. Randomized incremental construction of Delaunay and Voronoi diagrams. *Algorithmica* **7**, 381–413 (1992).

80. Chiang, W.-L. *et al.* Cluster-GCN: An Efficient Algorithm for Training Deep and Large Graph Convolutional Networks. in *KDD'19* (2019). doi:10.1145/3292500.3330925.

81. Hamilton, W. L., Ying, R. & Leskovec, J. Inductive Representation Learning on Large Graphs. in (2017).

82. Peter Bankhead. QuPath: Open source software for digital pathology image analysis | Scientific Reports. https://www.nature.com/articles/s41598-017-17204-5 (2017).

83. Labelbox. Labelbox. https://labelbox.com (2022).

84. McHugh, M. L. Interrater reliability: the kappa statistic. *Biochem. Medica* **22**, 276–282 (2012).

85. Paszke, A. *et al.* PyTorch: An Imperative Style, High-Performance Deep Learning Library. Preprint at https://doi.org/10.48550/arXiv.1912.01703 (2019).

86. Fey, M. & Lenssen, J. E. Fast Graph Representation Learning with PyTorch Geometric. *ArXiv190302428 Cs Stat* (2019).