HHS Public Access

Author manuscript

Mod Pathol. Author manuscript; available in PMC 2024 June 07.

Published in final edited form as:

Mod Pathol. 2023 August ; 36(8): 100196. doi:10.1016/j.modpat.2023.100196.

A Deep Learning Approach for Histology-Based Nucleus Segmentation and Tumor Microenvironment Characterization

Ruichen Rong[a], **Hudanyun Sheng**[a], **Kevin W. Jin**[a], **Fangjiang Wu**[a], **Danni Luo**[a], **Zhuoyu Wen**[a], **Chen Tang**[a], **Donghan M. Yang**[a], **Liwei Jia**[b], **Mohamed Amgad**[c], **Lee A.D. Cooper**[c], **Yang Xie**[a,d,e], **Xiaowei Zhan**[a,d,f,*], **Shidan Wang**[a,*], **Guanghua Xiao**[a,d,e,*]

[a]Quantitative Biomedical Research Center, Peter O'Donnell Jr. School of Public Health, UT Southwestern Medical Center, Dallas, Texas

[b]Department of Pathology, UT Southwestern Medical Center, Dallas, Texas

[c]Department of Pathology, Northwestern University Feinberg School of Medicine, Chicago, Illinois

[d]Department of Bioinformatics, UT Southwestern Medical Center, Dallas, Texas

[e]Simmons Comprehensive Cancer Center, UT Southwestern Medical Center, Dallas, Texas

[f]Center for the Genetics of Host Defense, UT Southwestern Medical Center, Dallas, Texas

Abstract

Microscopic examination of pathology slides is essential to disease diagnosis and biomedical research. However, traditional manual examination of tissue slides is laborious and subjective. Tumor whole-slide image (WSI) scanning is becoming part of routine clinical procedures and produces massive data that capture tumor histologic details at high resolution. Furthermore, the rapid development of deep learning algorithms has significantly increased the efficiency and accuracy of pathology image analysis. In light of this progress, digital pathology is fast becoming a powerful tool to assist pathologists. Studying tumor tissue and its surrounding microenvironment provides critical insight into tumor initiation, progression, metastasis, and potential therapeutic targets. Nucleus segmentation and classification are critical to pathology image analysis,

*Corresponding authors. Xiaowei.Zhan@utsouthwestern.edu (X. Zhan), Shidan. Wang@utsouthwestern.edu (S. Wang), Guanghua.Xiao@utsouthwestern.edu (G. Xiao).

Author Contributions

R.R., S.W., X.Z., and G.X. conceived and designed the study. R.R., H.S., S.W., and X.Z. developed methodology and performed implementation. R.R., K.J., X.Z., S.W., D.Y., and Y.X. performed the writing, review, and revision of the paper. R.R., H.S., Z.W., and C.T. provided statistical analysis and interpretation of data. F.W. and D.L. provided technical support and web tool development. Y.Z., L.C., L.J., and M.A. provided pathologic insights and additional comments. All authors read and approved the final paper.

Ethics Approval and Consent to Participate

All patient data to be used in this study are de-identified. This research does not involve collection of patient samples. Therefore, this research is not considered to be human subjects research. This is confirmed by the NIH Human Subjects Decision Tool (grants.nih.gov/policy/humansubjects/research.htm), which explicitly states "because the specimens or data were not collected specifically for your study and no one on your study team has access to the subject identifiers linked to the specimens or data, your study is not considered human subjects research.

Supplementary Material

The online version contains supplementary material available at https://doi.org/10.1016/j.modpat.2023.100196

Declaration of Competing Interest

The authors report no relevant conflicts of interest.

especially in characterizing and quantifying the tumor microenvironment (TME). Computational algorithms have been developed for nucleus segmentation and TME quantification within image patches. However, existing algorithms are computationally intensive and time consuming for WSI analysis. This study presents Histology-based Detection using Yolo (HD-Yolo), a new method that significantly accelerates nucleus segmentation and TME quantification. We demonstrate that HD-Yolo outperforms existing WSI analysis methods in nucleus detection, classification accuracy, and computation time. We validated the advantages of the system on 3 different tissue types: lung cancer, liver cancer, and breast cancer. For breast cancer, nucleus features by HD-Yolo were more prognostically significant than both the estrogen receptor status by immunohistochemistry and the progesterone receptor status by immunohistochemistry. The WSI analysis pipeline and a real-time nucleus segmentation viewer are available at https://github.com/impromptuRong/hd_wsi.

Keywords

digital pathology; nucleus segmentation; whole-slide imaging; Yolo

Introduction

Histopathologic examination of tissues is the cornerstone of disease diagnosis and prognosis. In this procedure, a pathologist examines cells, their spatial organization and tissue structure, under a microscope, seeking hints of disease and its effect on the tissue. It requires experienced pathologists to identify and interpret subtle morphologic patterns in histopathology slides. Manual examination of pathology slides is laborious and time consuming; moreover, the results are subjective among hospitals and individuals. With advances in imaging technology, whole-slide image (WSI) scanning of tissue slides is becoming a routine clinical procedure. Advanced whole-slide scanners are capable of rapidly producing massive numbers of WSIs that capture histopathologic details in high resolution. As a result, digital pathology has seen increasing adoption in clinical practice and diagnosis. Recently, deep learning algorithms have been developed and used for pathology image analysis.[1-3] The advanced algorithms have been applied in primary diagnosis,[4-6] prognosis studies,[7,8] and association analyses between pathology image and genomic data.[9]

A tumor is a mass of tissue with complex structures consisting of cancer cells and surrounding nonmalignant cells, which form the tumor microenvironment (TME). Studying and characterizing TME can offer valuable insights into the initiation, progression, and metastasis of tumors, as well as potential therapeutic targets. Nucleus segmentation and classification are key steps in digital pathology image analysis. Several algorithms have been developed to precisely identify nuclei in pathology image analysis, which enable researchers to characterize and quantify TME by extracting TME-related imaging features and associating these features with patient outcomes and genomic information. For example, Wang et al[10] proposed the Histology-based Digital (HD) Staining algorithm for nucleus localization, classification, and masking. HD-Staining relies on the Mask R-CNN[11] instance segmentation algorithm. Mask R-CNN is constructed from the following components: (1) a deep convolutional neural network (CNN) backbone with a feature pyramid network (FPN)[12] to extract imaging features from the raw image, (2) a region proposal network[13]

to localize potential objects, (3) a detection head to classify detected objects into different types, and (4) a mask head to segment the boundary of detected objects. Mask R-CNN processes object localization and classification separately, and therefore, it is considered a 2-stage object detection algorithm. Alternatively, algorithms that merge the second and third components are considered 1-stage object detection algorithms. In earlier years, 1-stage object detection algorithms had the advantage in computational efficiency, whereas 2-stage object detection algorithms achieved better accuracy and coverage by sacrificing speed, especially on small objects. Using Mask R-CNN as its main nucleus segmentation algorithm, HD-Staining takes hours to analyze a single slide and is inefficient at analyzing large-scale data sets.

To reduce the computational cost of nucleus segmentation, Graham et al[14] proposed HoVer-Net, an approach that directly applied 3 distinct branches atop image features extracted by a VGG (Visual Geometry Group) backbone: a nucleus classification branch, a nucleus pixel branch, and a HoVer branch. HoVer-Net relied on energy-based post-processing to generate nucleus masks from the outputs of the nucleus pixel branch and HoVer branch. Owing to its network design and memory cost, HoVer-Net accepts relatively small patches (80 × 80 pixels) as input and takes significant computational time for its post-processing masks.

Both HD-Staining and HoVer-Net algorithms perform well in the analysis of pathology image patches. However, their computational requirements are prohibitively large for WSIs, which greatly limits their application to WSI analysis. Recent developments in deep learning brought forth more efficient object detection algorithms applicable to varying speed and accuracy requirements. For example, Cascade R-CNN[15] provides high-quality 2-stage detection for dense and occluded objects. EfficientDet[16] improves detection efficiency and accuracy under resource constraints by automatically selecting network designs through a compound scaling approach and achieves the highest performance in comparison to other networks across multiple benchmarks. FCOS[17] simplifies 1-stage detection by completely avoiding anchor box operations. DETR[18] and deformable-DETR[19] are end-to-end object detectors by directly applying a transformer encoder—decoder on top of feature maps. Accounting for algorithm stability and efficiency, the most widely used 1-stage detection algorithms are that of the Yolo family.[20] With advanced training and network design, (Scaled-)Yolov4,[21,22] YoloR,[23] YoloX,[24] Yolov6,[25] and Yolov7[26] enable 1-stage detection algorithms to achieve comparable performance to that of 2-stage algorithms.

In this study, we developed Histology-based Detection using Yolo (HD-Yolo), an algorithm for nucleus detection, segmentation, classification, and TME-related feature extraction for WSI analysis. HD-Yolo enhances computational efficiency and enriches slide-level TME features. HD-Yolo provides 3 key contributions: (1) HD-Yolo enables simultaneous detection and segmentation by including a segmentation module within the traditional Yolo architecture. (2) HD-Yolo significantly increases the computation speed and reduces the requirement of computational resources for WSI analysis. (3) HD-Yolo uses a density-based TME feature extraction module to summarize WSI-level and region of interest (ROI)—level features. We demonstrated the performance of HD-Yolo by applying it to different tissue types: lung cancer, liver cancer, and breast cancer. We further applied HD-Yolo to The Cancer Genome Atlas—Breast Cancer (TCGA-BRCA) data set[27] and illustrated the

prognostic value of image features extracted by HD-Yolo through survival analysis. Finally, to demonstrate its usage, we developed a web front end that implements HD-Yolo and displays WSI results.

Materials and Methods

Data Sets

Lung Adenocarcinoma Data Set—The lung adenocarcinoma data set curated in the HD-Staining paper[10] contains 127 patches (500 × 500 pixels) extracted from 39 pathologic ROIs in the National Lung Screening Trial data set.[28] Image patches were split into training, validation, and testing sets based on slide ID. Specifically, 105 patches from 29 slides were assigned to training, 12 patches from another 5 slides were assigned to validation, and the other 10 patches from the remaining 5 slides were reserved for testing. Nuclei in these image patches were manually labeled and segmented by a board-certified pathologist (L.J.) into 6 different categories: tumor nuclei, stromal nuclei, lymphocyte nuclei, macrophage nuclei, red blood cells, and karyorrhexis. More than 12,000 cell nuclei were included in the training set (tumor nuclei 24.1%, stromal nuclei 23.9%, lymphocytes 29.5%, red blood cells 5.8%, and others 16.7%), whereas 1227 and 1086 nuclei were included in the validation and testing sets, respectively.

Liver Tissue Data Set—The liver tissue data set is curated from 76 hepatic hematoxylin and eosin (H&E) slides (×40 magnification) in liver tissues. These slides consist of 2 normal mouse liver slides, 9 normal human liver slides, 2 cirrhotic human liver slides without hepatocellular carcinoma, and 63 cirrhotic human liver slides with hepatocellular carcinoma. Fifty-one image patches (500 × 500 pixels) in nonmalignant regions were randomly selected, and nuclei in these patches were manually labeled by pathologist (L.J.) into 4 categories: hepatocyte nuclei, stroma nuclei, lymphocyte nuclei, and red blood cells. The 51 image patches were split into 35 for training, 8 for validation, and 8 for testing.

Breast Cancer Data Set—The publicly available NuCLS data set[29] consists of many image patches extracted from breast cancer images from the TCGA Data set.[27] Image patches were split into single-rater and multi-rater data sets (inferred P-truth testing data set). Nuclei within the patches were annotated through the collaborative effort of pathologists, pathology residents, and medical students. The corrected single-rater data set was collected with high-quality annotations from a team of 25 nonpathologists and further corrected and approved by 1 of the 7 experienced pathologists. The single-rater data set was split into 5 different folds (fold1-fold5) for training and validating nucleus detection and segmentation algorithms. The multi-rater data set is a relatively smaller cohort validated by multiple pathologists and is, therefore, suitable as an independent testing data set. The corrected single-rater data set contains 1744 image patches, 54,916 annotations with class labels, and 27,976 unlabeled annotations. The multi-rater evaluation data set contains 53 image patches, 1203 annotations with class labels, and 150 unlabeled annotations. The nuclei were assigned to 20 subcategories and 4 superclasses: tumor, stromal, stromal tumor-infiltrating lymphocytes (sTILs), and others. All models compared in this study were trained and validated on fold1 in the corrected single-rater data set and further evaluated on the

multirater data set. Here, we report superclass performances for ease of comparison with other models and the statistics reported in the original papers.

The Cancer Genome Atlas—Breast Cancer Data Set—We analyzed the prognosis of breast carcinoma based on the H&E slides and patient clinical information available in the TCGA-BRCA data set. Among all 1061 patients with pathology images, 653 infiltrating ductal carcinoma (IDC) and 169 infiltrating lobular carcinoma (ILC) female patients with confirmed tumor stage information were used in this study. Patients were split into 80% training (658 patients) and 20% testing (164 patients) stratified by histologic types. Supplementary Table S1 summarizes the patient demographic and clinical information in the training and testing data sets.

Methodology in Histology-Based Detection Using Yolo Framework

Figure 1 illustrates the overall workflow of the HD-Yolo system. For each slide, we first divided the whole tissue region into small patches (640 × 640 pixels at ×40 magnification). We used a Yolo-based object segmentation algorithm (Fig. 1A) to quickly locate nucleus coordinates, identify nucleus type, and extract shape morphologic features. We then summarized the above information into spatial TME features using the feature extraction pipeline (Fig. 1B). The feature extraction pipeline focuses on selected image patches in the ROI and analyzes the entire slide for a comprehensive characterization of the TME.

Histology-Based Detection Using Yolo Nucleus Segmentation Model Architecture

The nucleus detection and segmentation model in the HD-Yolo system were constructed under the Yolo architecture with a customized design, as illustrated in Figure 1A. We used Yolov5's CSPDarkNet as the backbone for image feature extraction.[22] CSPDarkNet uses Cross Stage Partial Network layers to partition and merge the original feature maps in DarkNet for cross-stage communication. We then applied the Bi-FPN[16] to blend backbone feature maps at different resolutions. We further used the YoloX decoupled detection head[24] to localize and classify objects. As 1-stage object detection algorithms, the Yolo family cannot directly provide object masks for detected objects; therefore, we built an additional multi-ROIAlign + fully convolutional network (FCN) head on top of the detection workflow. For the nucleus detection task, we observed that the YoloX anchor-free decoupled detection head achieved better accuracy than the merged head used in Yolov3 to Yolov5. The decoupled head was capable of further splitting nucleus location and type information for representative analysis. For the segmentation task, our segmentation balanced accuracy and speed. The Bi-FPN head provides rich information for both nucleus masking and background segmentation. Compared with the Mask R-CNN mask head, we reduced the number of convolutional layers and channels to greatly increase the computational speed and memory cost. Compared with prototype-based method deployed in YoloACT,[30] YoloACT++,[31] and Yolov5 segmentation, our approach can generate more accurate masks for occluded nuclei with competitive speed. With the mask head added, HD-Yolo is not strictly considered a 1-stage object detection algorithm. However, its overall computational speed is competitive with that of the latest Yolo variations and is still much faster than existing nucleus segmentation algorithms.

Whole-Slide Image Nucleus Density-Based Tumor Microenvironment Feature Extraction

Owing to computational speed limitations, most existing TME analysis systems are deployed on a limited number of randomly selected patches when analyzing large-scale high-resolution slides. For instance, the HD-Staining pipeline characterizes nucleus spatial organization by building Delaunay triangle graphs and summarizes nuclei connections in 100 patches (1024 × 1024 pixels) randomly sampled from the ROI. This approach relies on human-labeled ROIs, and the results would be potentially biased according to the size and number of available patches inside the ROI. When ROIs are large, important subregions may be ignored from calculation. In contrast, when ROIs are small, overlapped subregions are likely to be analyzed multiple times. To alleviate the overlapping issue, one may sample a small number (eg, 10) of larger patches (2048 × 2048 pixels) to include more connections in a single image, but this increases variance during summarizationSampling a higher number (eg, 100) of smaller patches (1024 × 1024 pixels) reduces between-patch variance but excludes connections at the patch border from the calculation. Our HD-Yolo system enables fast nucleus segmentation over whole slides and thus makes summarizing spatial properties globally feasible. We refined the procedure of Delaunay triangulation by automatically selecting the ROI based on nucleus densities without human labeling. In addition to ROI-based TME features, we expanded several slide-level TME features by analyzing nucleus distribution and densities over all nuclei in the slide (Fig. 1B).

The following comprises HD-Yolo's TME feature extraction process. First, the nucleus segmentation algorithm is run on whole slides, and the coordinates, types, and sizes of all detected nuclei are recorded. Each type of nuclei is then allocated into a 2D point cloud, and kernel smoothing[32] is applied based on nucleus size to generate a density map. These density maps reflect the distribution of different types of nuclei in the slide. To extract slide-level TME features, the following intensity features are calculated for each pair of nuclei *type_i* and *type_j* based on their density maps *map_i* and *map_j*: (1) the dot product between *map_i* and *map_j*, which summarizes the overall interactions between *type_i* and *type_j*; (2) the projection from *map_i* to *map_j*, which represents the level of *type_j* influenced by *type_i* and vice versa; and (3) the cosine similarity between *map_i* and *map_j*, which represents the interaction intensity between *type_i* and *type_j*. These density features are irrelevant to the size and number of patches sampled and are robust to ROI size variation. To extract ROI-based TME features, the above analysis is used to automatically decide ROIs and build a Delaunay triangle graph in each region. By default, regions with high tumor densities and regions with high interaction intensities are potential ROIs. Delaunay triangle triplets are calculated in each region, and the edge probabilities for each pair of nuclei *type_i* and *type_j* are averaged over all regions.

Histology-Based Detection Using Yolo Performance Evaluation Criteria

Evaluation of Nucleus Detection and Segmentation Performance—We compared HD-Yolo with existing models using the following statistics:

- Detection coverage: the percentage of ground truth nuclei detected by the model
- Nucleus classification accuracy: the percentage of ground truth nuclei that was further correctly classified by the model

- Matthews correlation coefficient (MCC): a score of the model's overall performance considering accuracy and coverage among different classes
- Median intersection over union (mIoU): the mIoU between the predicted mask and the original mask measures the segmentation similarity between detected nuclei and ground truth nuclei
- Precision: the percentage of true nuclei among detected nuclei
- Recall rate: how many ground truth nuclei are detected
- F1 score: the harmonic mean of precision and recall
- Time per image: the mean inference speed for the whole data set on a Tesla V100 GPU with 32 GB of memory.

Coverage, accuracy, MCC, precision, recall, and F1 score were evaluated at thresholding, with the intersection over union of the ground truth box and prediction box >0.5. For precision, recall, and F1 score, we eliminated the unbalancing between each class by calculating the class-specific statistics and then averaging the scores over all classes. For time per image, we recorded the total computational time from when the image was loaded into the model to when the results of the whole data set were exported and then averaged over the number of images. More specifically, we included the postprocessing time of each model in the calculation, whereas file reading, plotting, and result exporting times were excluded. We selected the highest batch size and number of processors to allow maximal computational resources (GPU, CPU, etc.) for each model.

Evaluation of Whole-Slide Image Analysis and Feature Extraction—We compared the WSI inference and feature extraction times of different algorithms when processing slides of different sizes from the TCGA-BRCA data set. The regularized Cox proportional hazard (CoxPH) model[33-35] was used to analyze prognosis hazard ratios, and the ElasticNet algorithm was used for feature selection. We applied 2 prognosis analyses separately on IDC patients alone and on all IDC and ILC patients. As there were 0 death events for ILC in the testing data set, we did not perform separate analyses for ILC. For each data set, we built multiple survival models with different sets of features: (1) clinical features only; (2) clinical features plus Delaunay graph-based TME features defined in HD-Staining; (3) clinical features plus density-based TME features as defined in the Whole-Slide Image Nucleus Density-Based Tumor Microenvironment Feature Extraction section; and (4) ElasticNet-selected features from all features listed above.

For each survival model, the C index is used to quantify the model's ability to correctly provide a reliable ranking of the survival times by calculating the concordance between predicted risk scores and overall survival. We first used a grid search strategy to select the best hyperparameters according to the 10-fold cross-validation C index and then applied the model with the highest cv C index to the testing data set. We report the cross-validation C index in the training data set and the C index (mean ± SE) in the testing data set to measure the performance of each survival model. Moreover, we have calculated the time-dependent receiver operation characteristic curve and determined the area under the curve (mean ±

SE) on the testing data set for the survival analysis. The feature coefficients, permutation importance, and *P* values are also provided to demonstrate the feature selection results.

Histology-Based Detection Using Yolo Implementation Details

HD-Yolo was developed in Python 3.9. The Yolo-based nucleus segmentation algorithm was modified from the official YoloX (https://github.com/MegEngine/YOLOX) and Yolov5 repositories (https://github.com/ultralytics/yolov5, commit: 7f5724ba4b3e42 1d4c9162742810c52248d06ecd). We modified the Bi-FPN backbone and YoloX head with our mask support and immigrated the model into the Yolov5 distributed data infrastructure. We further reimplemented the training, validation, and testing pipeline to support intensive data augmentation, joint detection, and segmentation training and customized the data with missing and unclassified labels/masks and whole-slide inference. Major packages used in HD-Yolo development and evaluation include PyTorch 1.10.1, torchvision 0.11.2, scikit-image 0.18.3, scikit-learn 1.0.2, and openslide 3.4.1. For training and inference, we distributed the data across 4 × 32 GB Tesla V100 GPUs and deployed the model on a Tesla V100 GPU. The regularized CoxPH model was built with the scikit-survival 0.17.2 package, and the permutation importance score was calculated with ELI5 0.11.0 in Python. The survival statistics C index and *P* value were calculated using the R packages *survival*, *glmnet*, and *survcomp* in R 3.6.1.

For comparison purposes, we implemented the following algorithms as described in their original publications and default settings in GitHub:

- HD-Staining: https://github.com/matterport/Mask_RCNN (Mask R-CNN with resnet50 backbone, commit: 3deaec5d902d16e1daf56b62d5971d428dc920bc)
- Mask R-CNN: We customized the PyTorch Mask R-CNN with EfficientNet-B3 backbone[36] and softNMS[37] for better accuracy and faster speed
- HoVer-Net: https://github.com/vqdang/hover_net
- Yolov7: https://github.com/WongKinYiu/yolov7
- Yolov6: https://github.com/meituan/YOLOv6
- ScaledYolov4: https://github.com/WongKinYiu/ScaledYOLOv4
- FCOS: https://github.com/Adelaide-AI-Group/FCOS
- Deformable-DETR: https://github.com/open-mmlab/mmdetection
- EfficientDet: https://github.com/rwightman/efficientdet-pytorch
- Cascade R-CNN: https://github.com/open-mmlab/mmdetection.

Results

Histology-Based Detection Using Yolo Nucleus Segmentation Performance

We compared HD-Yolo's performance on the lung cancer data set with that of 3 existing nucleus segmentation algorithms: HD-Staining, PyTorch Mask R-CNN, and HoVer-Net. All models were evaluated using 500 × 500 image patches at ×40 magnification. Nuclei

were detected, segmented, and classified into 6 categories: tumor nuclei, stromal nuclei, lymphocyte nuclei, macrophage nuclei, red blood cells, and karyorrhexis. The performance and speed of each model on the testing and validation data sets are listed in Table 1[38] and Supplementary Table S2.

We observed that HD-Yolo achieved superior performance and speed among all models: HD-Yolo has ~5% greater coverage on ground truth nuclei than HD-Staining and HoVer-Net, HD-Yolo achieves ~6% higher accuracy than HD-Staining and ~12% higher accuracy than HoVer-Net, and HD-Yolo is 15 times faster than HD-Staining (0.01 seconds vs 0.15 seconds) and 100 times faster than HoVer-Net (0.01 seconds vs 1 second).

For the task of nucleus segmentation, HD-Yolo generated smooth, round masks for different types of nuclei and provided accurate masks for occluded nuclei. HD-Yolo's mIoU is competitive with that of HD-Staining and Mask R-CNN and is significantly higher than that of HoVer-Net and Yolov5 segmentation. As shown in Figure 2, we observed that HoVer-Net failed to segment occluded nuclei and tended to split 1 large nucleus into several smaller ones. Yolov5 segmentation tends to generate rectangle-shaped nuclei for occluded objects. We ascribe the success of our segmentation head to the following reasons: (1) Detection features are generally distinct from segmentation features, especially for 1-stage detection algorithms, and Bi-FPN provides important intrinsic image features for precise segmentation. (2) Nucleus morphologic features are relatively similar across different types, and a light-weight FCN can provide satisfactory masks at a considerable low computational cost. (3) The ROIAlign + FCN-based mask heads (used by HD-Yolo, HD-Staining, and Mask R-CNN) are more robust than heuristic post-processing (HoVer-Net) and prototype-based mask head (YoloACT, Yolov5 segmentation) on irregularly shaped nuclei and occluded nuclei.

We further used transfer learning to build a liver nucleus segmentation model from the above lung cancer HD-Yolo model (Fig. 3A). The liver cancer data set contained only 35 training, 8 validation, and 8 testing patches, which is relatively small compared with the lung cancer study. The model achieved 0.9287 coverage and 0.8437 accuracy on the testing data set after only 10 epochs of fine-tuning (Fig. 3B), demonstrating HD-Yolo's generalizability.

Histology-Based Detection Using Yolo Object Detection Performance

In addition, we compared HD-Yolo detection performance with that of advanced object detection algorithms on the large-scale NuCLS benchmark data set. The NuCLS data set contains missing labels and lacks complete mask annotations; moreover, not all models included in this comparison were able to provide masks without extra modification. Therefore, we focused only on detection performance based on 3 predefined superclasses: tumor, stromal, and sTILs. As shown in Table 2, HD-Yolo achieved similar performance to that of a variety of advanced object detection algorithms. We observed that 1-stage algorithms (HD-Yolo, Yolov6, and Yolov7) and 2-stage algorithms (EfficientDet and Cascade R-CNN) achieved the highest F1 score. However, HD-Yolo, Yolov6, and Yolov7 outperformed other models in computational speed. Surprisingly, FCOS and deformable-DETR maintained comparable precision but had significantly lower coverage compared with other methods. This could be because of FCOS and deformable-DETR's use of heat

map and transformer-based approaches, which are not suitable for cases where the same type of nuclei at different sizes are densely overlapped. Overall, we believe that HD-Yolo can achieve state-of-the-art performance compared with currently available object detection algorithms and provide reliable breast cancer WSI analysis results.

Histology-Based Detection Using Yolo Whole-Slide Image Analysis

HD-Yolo Whole-Slide Image Inference With Single and Ensembled Models—

Ensemble modeling is a common method for improving overall inference performance by combining multiple models without altering the network structure. To build a more accurate breast cancer model, we ensembled 5 independent nucleus segmentation models trained and validated on 5 predefined splits (fold1-fold5) in the NuCLS corrected single-rater data set. During inference, the input images went through 5 different models, and duplicated detections were filtered by nonmaximum suppression. We further halved the inference speed without influencing performance by switching from singleprecision (float32) models to half-precision (float16) models. Table 3 shows the performance and speed of the ensembled model and every independent model on the inferred P-truth testing data set. The ensemble model achieved even higher performance, with an ~5% higher detection rate (coverage), ~7% higher detection accuracy, ~6% higher recall, and ~7% higher MCC compared with the single models. Therefore, we believe that both the single models and the ensembled model can be generalized to large-scale breast cancer whole-slide analysis for TME feature extraction; single models may be used for fast screening, whereas the ensembled model can provide more accurate nucleus detection results. Performance for each superclass (tumor, stromal, and sTILs) and confusion matrices are summarized in Supplementary Tables S3 and S4.

We analyzed all H&E slides in the TCGA-BRCA data set and compared inference speeds between HD-Yolo single, HD-Yolo ensemble, HD-Staining, and HoVer-Net. For HD-Yolo and HD-Staining, we roughly extracted the tissue region from each slide by thresholding and scanned the slides with a 512 × 512 window. For HoVer-Net, we used the default tissue extraction and patch preparation pipeline as defined in the original repository. For HD-Staining and HoVer-Net, we used the maximum possible batch size that fit in a 32-GB memory GPU and the maximum amount of CPU for postprocessing. For HD-Yolo, we limited GPU memory to <16 GB (for a 512 × 512 window with a batch size of 32) and used at most 32 CPU threads (16 core) to mimic a small server setup with balanced CPU and GPU resources. An example of HD-Yolo's whole-slide inference results can be found in Supplementary Figure S1. Table 4 summarizes whole-slide inference speeds for all models on the TCGA-BRCA data set based on tissue sizes under different quantiles: minimum, median, 95% quantile, and maximum. The HD-Yolo single model took only 1/30 to 1/25 of the time used by HD-Staining and 1/130 to 1/50 of the time used by HoVer-Net. The more accurate ensemble HD-Yolo model took double the processing time of the single version. Accounting for the substantial processing time expended when exporting the morphologic results of tens of thousands of nuclei, HD-Yolo's overall speed is about 10 to 15 times faster than HD-Staining and 50 times faster than HoVer-Net. It is noteworthy that in this comparison, HD-Yolo uses only half of the computational resources used by HD-Staining and HoVer-Net but is already 50 to 100 times faster. We expect even greater improvement on

large slides with more GPU memory and CPU cores by increasing image patch size, batch size, and industrial/organizational workers. As the majority of slides (>95% quantile) took <5 minutes to process with the single model and ~10 minutes with the ensemble model, the whole data set (1061 slides) took about 12 hours with the single model and 1 day with the ensemble model on a server with 4 GPUs. The same task would take 1 to 2 weeks with HD-Staining and longer with HoVer-Net.

The Cancer Genome Atlas Breast Cancer Survival Analysis Based on Clinical and Nucleus Spatial Features

We extracted TME features based on the nucleus segmentation results from the fast HD-Yolo single model computed in the Histology-Based Detection Using Yolo Whole-Slide Image Analysis section and built a CoxPH model for breast cancer prognosis. For each slide, we constructed a spatial density map based on the nucleus type, location, and shape morphologic features of nuclei belonging to the 3 superclasses: tumor, stromal, and sTILs. We include in the analysis the 15 Delaunay graph-based interaction features defined in HD-Staining and the 12 nucleus density-based TME features in HD-Yolo defined in the Whole-Slide Image Nucleus Density-Based Tumor Microenvironment Feature Extraction section. In addition to the image-based TME features, the following clinical features are included in the analysis: age at diagnosis, American Joint Committee on Cancer pathologic tumor stage (American Joint Committee on Cancer tumor stage), estrogen receptor status by immunohistochemistry (IHC), progesterone receptor status by IHC, and human epidermal growth factor receptor 2 by IHC. Detailed information regarding the TME and clinical features can be found in Supplementary Table S5.

We applied the regularized CoxPH model to analyze prognosis hazard ratios and used the ElasticNet algorithm for feature selection. The cross-validation C index, the testing C index (mean ± SE), and the testing time-dependent area under the curve (mean ± SE) for different models and feature sets can be found in Figure 4A. We observed that, despite the clinical features of age at diagnosis and tumor stage, tumor-infiltrating lymphocytes (TIL)-related TME features defined by HD-Yolo are highly correlated with survival outcomes (Supplementary Table S6 provides feature importance scores and *P* values). *l_t.proj.prob* has the highest negative coefficient, whereas *t_l.proj.prob* has the second highest positive coefficient other than age (Fig. 4B). These 2 TIL-related TME features are more prognostically significant and have a stronger predictive value than the estrogen receptor status by IHC, progesterone receptor status by IHC, and human epidermal growth factor receptor 2 by IHC. The *l_t.proj.prob* ($P < .05$) feature represents the degree to which tumor tissue has been invaded by sTILs. A high *l_t.proj.prob* indicates that large numbers of tumors in tissue regions have been invaded by sTILs and are strongly interacting with the lymphocytes. This is consistent with the prognostic model that an increase in *l_t.proj.prob* will likely reduce the hazard ratio and lead to better survival outcomes. In contrast, the *t_l.proj.prob* ($P < .01$) feature represents the degree to which sTILs are surrounded by tumor cells. A high *t_l.proj.prob* indicates that lymphocytes have been overwhelmed by surrounding tumor nuclei and thus cannot effectively inhibit tumor tissue growth. This is consistent with the prognostic model that an increase in *t_l.proj.prob* will likely increase the hazard ratio and lead to worse survival outcomes.

Discussion

In this paper, we proposed HD-Yolo, a system for fast, accurate nucleus segmentation and rich TME feature extraction. Our comparison studies show that the HD-Yolo algorithm outperformed existing algorithms (Mask R-CNN, HoVer-Net, etc.) in nucleus detection and segmentation accuracy. HD-Yolo also provided high-quality masks without sacrificing efficiency in contrast to advanced object detection algorithms. With its fast nucleus detection algorithm, HD-Yolo significantly accelerated WSI analysis while using less computational resources. In addition, the TME features summarized by HD-Yolo's nucleus density-based method are highly correlated with breast cancer patient prognosis (prognosis was not assessed for lung or liver) and provide meaningful explanations regarding tumor progression and survival. The efficient and accurate HD-Yolo pipeline can be useful for a variety of tasks in pathology imaging analysis. In summary, we expect HD-Yolo to be a powerful tool that will facilitate the analysis of digital pathology and provide meaningful biological insights.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

This work was supported by the National Institutes of Health (R01GM140012, R01GM141519, R01DE030656, U01CA249245, R35GM136375, P50CA070907, and 2P30CA142543) and the Cancer Prevention and Research Institute of Texas (CPRIT RP180805, RP230330 and RP190107). The funding bodies had no role in the design, collection, analysis, or interpretation of data in this study.

Data Availability

For lung cancer analysis, the pathology patches were extracted from the National Lung Screening Trial data set (https://cdas.cancer.gov/nlst/). For breast cancer, pathology patches and annotations that were used for algorithm training, validation, and comparison are available in the online NuCLS portal (https://sites.google.com/view/nucls/). The whole-slide imaging support the findings of the survival study is available at The Cancer Genome Atlas —Breast Cancer (https://www.cancer.gov/types/breast).

References

1. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J Pathol Inform. 2016;7:29. [PubMed: 27563488]
2. Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60–88. [PubMed: 28778026]
3. Wang S, Yang DM, Rong R, Zhan X, Xiao G. Pathology image analysis using segmentation deep learning algorithms. Am J Pathol. 2019;189(9): 1686–1698. [PubMed: 31199919]
4. Wang D, Khosla A, Gargeya R, Irshad H, Beck AH. Deep learning for identifying metastatic breast cancer. Preprint. Posted online June. 2016;18. ArXiv. doi: arXiv:1606.05718.
5. Folmsbee J, Liu X, Brandwein-Weber M, Doyle S. Active deep learning: improved training efficiency of convolutional neural networks for tissue classification in oral cavity cancer. In: Proceedings. IEEE International Symposium on Biomedical Imaging. IEEE Computer Society; 2018: 770–773.

6. Amgad M, Elfandy H, Hussein H, et al. Structured crowdsourcing enables convolutional segmentation of histology images. Bioinformatics. 2019;35(18):3461–3467. [PubMed: 30726865]
7. Tabibu S, Vinod P, Jawahar C. Pan-renal cell carcinoma classification and survival prediction from histopathology images using deep learning. Sci Rep. 2019;9(1), 10509. [PubMed: 31324828]
8. Cheng J, Mo X, Wang X, Parwani A, Feng Q, Huang K. Identification of topological features in renal tumor microenvironment associated with patient survival. Bioinformatics. 2018;34(6):1024–1030. [PubMed: 29136101]
9. Ren M, Zhang S, Zhang Q, Ma S. HeteroGGM: an R package for Gaussian graphical model-based heterogeneity analysis. Bioinformatics. 2021;37(18):3073–3074. [PubMed: 33638346]
10. Wang S, Rong R, Yang DM, et al. Computational staining of pathology images to study the tumor microenvironment in lung cancer. Cancer Res. 2020;80(10):2056–2066. [PubMed: 31915129]
11. He K, Gkioxari G, Dollár P, Girshick R. Mask r-cnn. In: Proceedings. IEEE International Conference on Computer Vision. IEEE Computer Society; 2017:2961–2969.
12. Lin TY, Dollár P, Girshick R, He K, Hariharan B, Belongie S. Feature pyramid networks for object detection. In: Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society; 2017:2117–2125.
13. Ren S, He K, Girshick R, Sun J. Faster r-cnn: towards real-time object detection with region proposal networks. Adv Neural Inf Process Syst. 2015;28:91–99.
14. Graham S, Vu QD, Raza SEA, et al. Hover-net: simultaneous segmentation and classification of nuclei in multi-tissue histology images. Med Image Anal. 2019;58, 101563. [PubMed: 31561183]
15. Cai Z, Vasconcelos N. Cascade r-cnn: delving into high quality object detection. In: Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society; 2018:6154–6162.
16. Tan M, Pang R, Le QV. Efficientdet: scalable and efficient object detection. In: Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society; 2020:10781–10790.
17. Tian Z, Shen C, Chen H, He T. Fcos: fully convolutional one-stage object detection. In: Proceedings. IEEE International Conference on Computer Vision. IEEE Computer Society; 2019:9627–9636.
18. Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S. End-to-end object detection with transformers. In: European Conference on Computer Vision. Springer-Verlag; 2020:213–229.
19. Zhu X, Su W, Lu L, Li B, Wang X, Dai J. Deformable detr: deformable transformers for end-to-end object detection. Posted online March 18 Preprint. 2021. ArXiv. https://doi.org/arXiv:2010.04159
20. Redmon J, Divvala S, Girshick R, Farhadi A. You only look once: unified, real-time object detection. In: Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society; 2016:779–788.
21. Bochkovskiy A, Wang CY, Liao HYM. Yolov4: optimal speed and accuracy of object detection. Posted online April 23 Preprint. 2020. ArXiv. https://doi.org/arXiv.2004.10934
22. Wang CY, Bochkovskiy A, Liao HYM. Scaled-yolov4: scaling cross stage partial network. In: Proceedings. IEEE Computer Society Conference on Computer Vision and Pattern Recognition. IEEE Computer Society; 2021:13029–13038.
23. Wang CY, Yeh IH, Liao HYM. You only learn one representation: unified network for multiple tasks. Posted online May 10 Preprint. 2021. ArXiv. https://doi.org/arXiv:2105.04206
24. Ge Z, Liu S, Wang F, Li Z, Sun J. Yolox: exceeding yolo series in 2021. Posted online August 6 Preprint. 2021. ArXiv. doi:arXiv:2107.08430.
25. Li CL, Li L, Jiang HJ, et al. YOLOv6: a single-stage object detection framework for industrial applications. Posted online September 7 Preprint. 2022. ArXiv. https://doi.org/arXiv:2209.02976
26. Wang CY, Bochkovskiy A, Liao HYM. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Posted online July 6 Preprint. 2022. ArXiv. https://doi.org/arXiv:2207.0269
27. Tomczak K, Czerwińska P, Wiznerowicz M. The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge. Contemp Oncol. 2015;19(1A):A68.

28. Kramer BS, Berg CD, Aberle DR, Prorok PC. Lung cancer screening with low-dose helical CT: results from the National Lung Screening Trial (NLST). J Med Screen. 2011;18(3):109–111. [PubMed: 22045816]
29. Amgad M, Atteya LA, Hussein H, et al. NuCLS: a scalable crowdsourcing approach and dataset for nucleus classification and segmentation in breast cancer. Gigascience. 2022;11:giac037. [PubMed: 35579553]
30. Bolya D, Zhou C, Xiao F, Lee YJ. Yolact: real-time instance segmentation. In: Proceedings. IEEE International Conference on Computer Vision. IEEE Computer Society; 2019:9157–9166.
31. Bolya D, Zhou C, Xiao F, Lee YJ. Yolact++: better real-time instance segmentation. IEEE Trans Pattern Anal Mach Intell. 2022;44:1108–1121. [PubMed: 32755851]
32. Wand MP, Jones MC. Kernel Smoothing. CRC Press; 1994.
33. Breslow NE. Analysis of survival data under the proportional hazards model. Int Stat Rev. 1975;43:45–57.
34. Tibshirani R. The lasso method for variable selection in the Cox model. Stat Med. 1997;16(4):385–395. [PubMed: 9044528]
35. Zou H, Hastie T. Regularization and variable selection via the elastic net. J R Stat Soc Series B Stat Methodol. 2005;67(2):301–320.
36. Tan M, Le Q. Efficientnet: rethinking model scaling for convolutional neural networks. In: Proceedings of the International Conference on Machine Learning. PMLR; 2019:6105–6114.
37. Bodla N, Singh B, Chellappa R, Davis LS. Soft-NMS—improving object detection with one line of code. In: Proceedings. IEEE International Conference on Computer Vision. IEEE Computer Society; 2017:5561–5569.
38. Gamper J, Alemi Koohbanani N, Benet K, Khuram A, Rajpoot N. Pannuke: an open pan-cancer histology dataset for nuclei instance segmentation and classification. In: Proceedings. European Congress on Digital Pathology. Springer; 2019:11–19.

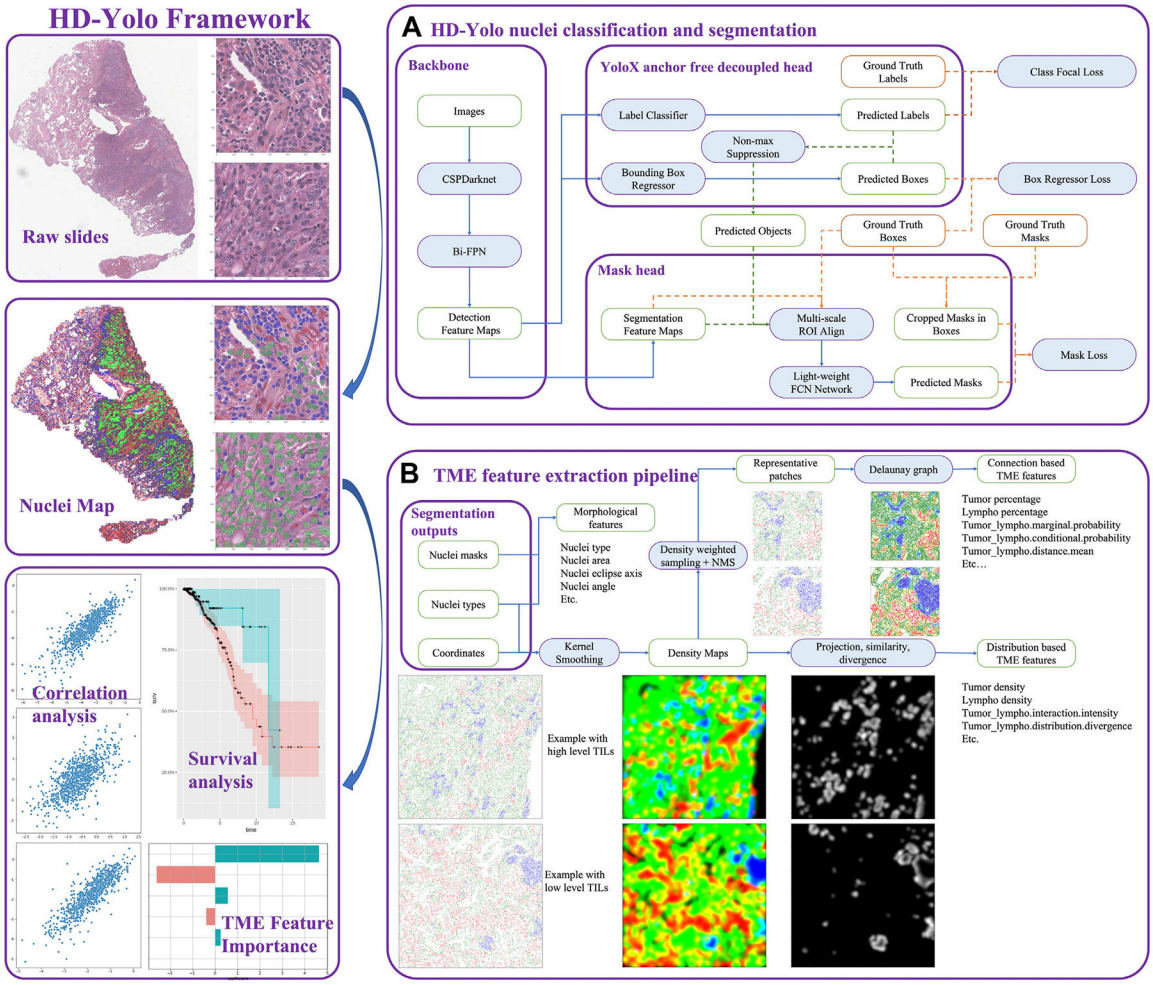


Figure 1.

HD-Yolo system workflow. (A) Nucleus detection and segmentation model in the HD-Yolo system. The model is constructed from 3 components: a Bi-FPN backbone, a YoloX detection head, and an ROIAlign + FCN mask head. The light blue boxes with purple borders are deep learning modules. The white boxes with green borders are intermediate outputs of the model. The solid blue lines are computational graphs used in both training and inference. The orange dotted lines are only processed in training, whereas the green dotted lines are for inference only. (B) TME feature extraction pipeline based on the nucleus density maps built from the nucleus detection results. The pipeline directly extracts whole-slide-level TME features (distributions, projections, similarity, and divergence) from the spatial densities of different types of nuclei. The pipeline also summarizes ROI-level TME features by automatically selecting ROIs from the above analysis. FCN, fully convolutional network; FPN, feature pyramid network; HD-Yolo, Histology-based Detection using Yolo; ROI, region of interest; TME, tumor microenvironment.

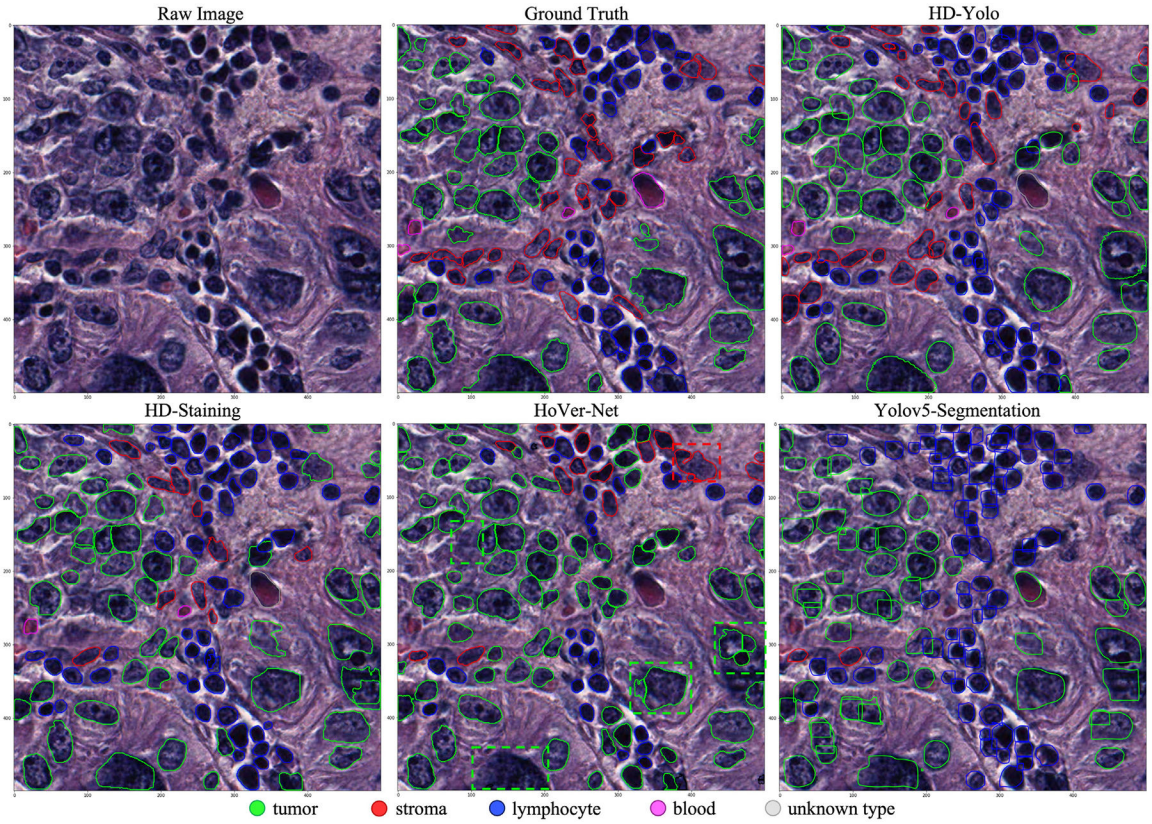
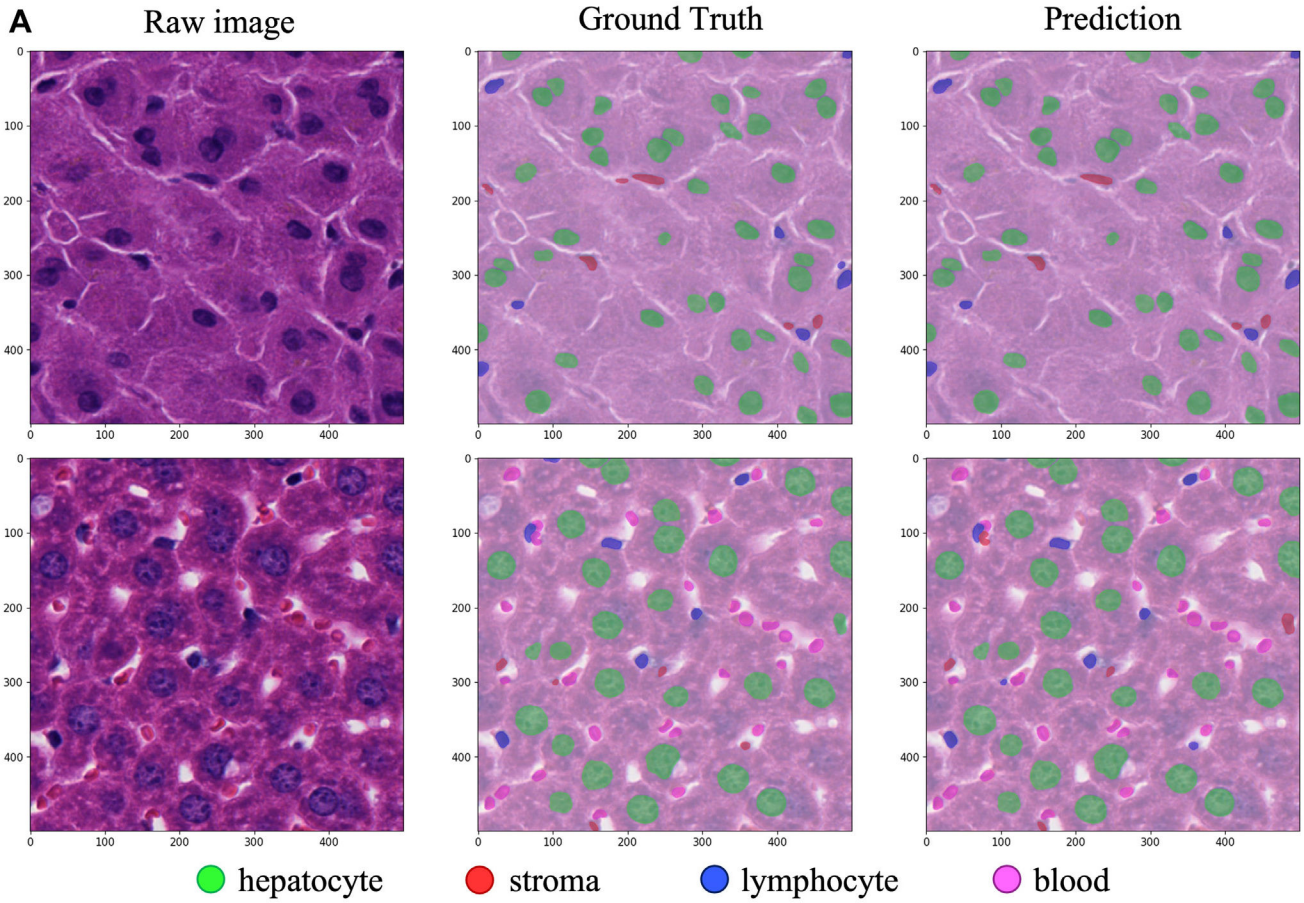


Figure 2.

Nucleus segmentation results for different models on a testing lung cancer image patch. HD-Yolo and HD-Staining generated round and smooth nuclei-shaped masks. HoVer-Net failed to segment occluded objects and a very large tumor nucleus (bottom left rectangle) and tended to split large nuclei into smaller pieces (examples are annotated with dashed bounding boxes). Yolov5 segmentation used prototype approach for masking. The algorithm generated rectangle boundaries for occluded nuclei. HD-Yolo, Histology-based Detection using Yolo.

A



B

Confusion Matrix

	hepatocyte	stroma	lymphocyte	others	missing
hepatocyte	267	5	0	0	14
stroma	12	36	7	0	7
lymphocyte	6	5	26	0	3
others	0	1	1	38	7

Normalized Confusion Matrix

	hepatocyte	stroma	lymphocyte	others	missing
hepatocyte	0.9336	0.0175	0.0000	0.0000	0.0490
stroma	0.1935	0.5806	0.1129	0.0000	0.1129
lymphocyte	0.1500	0.1250	0.6500	0.0000	0.0750
others	0.0000	0.0213	0.0213	0.8085	0.1489

Figure 3.

Liver HD-Yolo nucleus segmentation model transferred from the lung cancer model. (A) Two examples in the testing data set. (B) Raw and normalized confusion matrix of the testing data set. HD-Yolo, Histology-based Detection using Yolo.

A C-index with different feature sets

	IDC only model			IDC + ILC model		
	Crossval C-index	Test C-index (mean ± se)	Test AUC (mean ± se)	Crossval C-index	Test C-index (mean ± se)	Test AUC (mean ± se)
Clinical feature only	0.7743	0.7208±0.0649	0.7724±0.0183	0.8056	0.7659±0.0503	0.7828±0.0191
Clinical + Delaunay TME	0.7498	0.7075±0.0684	0.7555±0.0156	0.7991	0.7775±0.0443	0.7931±0.0112
Clinical + Density TME	0.7771	0.7878±0.0445	0.8107±0.0144	0.8081	0.8266±0.0323	0.8497±0.0162
ElasticNet Selected	0.7750	0.7514±0.0550	0.8017±0.0186	0.8180	0.8251±0.0363	0.8404±0.0131

B Important feature coefficients

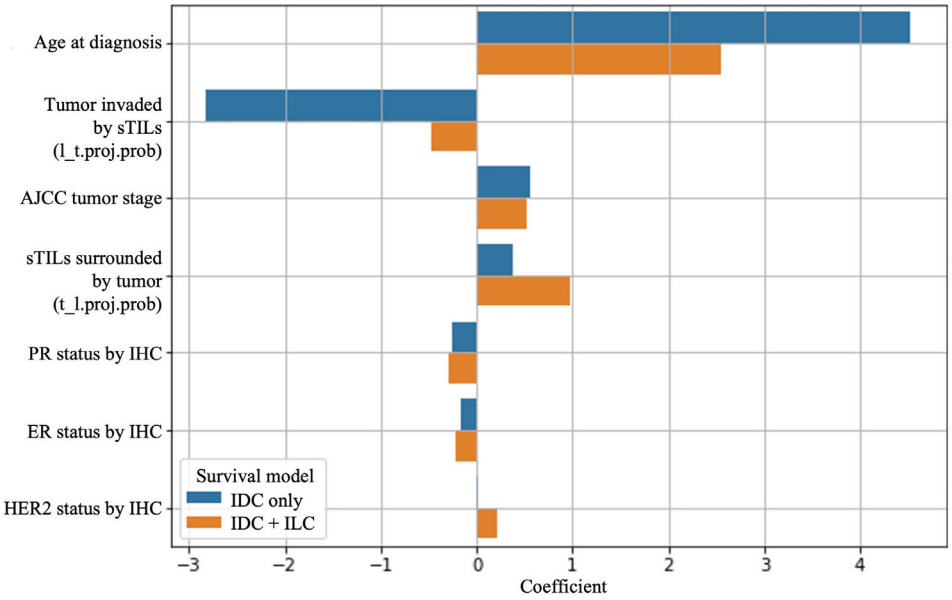


Figure 4.

CoxPH survival analysis for breast cancer subtypes. The study evaluates independent breast cancer survival models that incorporate clinical features and TME features for both IDC-only and IDC + ILC patients. (A) The top panel reports the cross-validation C index, testing C index (mean ± SE), and testing area under the curve (mean ± SE) under different data sets and feature sets. Clinical features include age at diagnosis, AJCC pathologic tumor stage, ER status by IHC, PR status by IHC, and HER2 status by IHC. Clinical + Delaunay TME uses clinical features with Delaunay graph-based features. Clinical + density TME includes clinical features with nucleus density-based features. ElasticNet Selected takes clinical features plus all image features with ElasticNet permutation importance > 0 and $P < .05$. (B) The bottom panel plots the coefficients of the selected features from ElasticNet. AJCC, American Joint Committee on Cancer; CoxPH, Cox proportional hazard; ER status by IHC, estrogen receptor status by immunohistochemistry; HER2 status by IHC, human epidermal growth factor receptor 2 by immunohistochemistry; IDC, Infiltrating ductal carcinoma; ILC, infiltrating lobular carcinoma; PR status by IHC, progesterone receptor status by immunohistochemistry; TME, tumor microenvironment.

Table 1

Comparison of nucleus detection performance and speed among different models on lung cancer testing data set

Testing	Coverage	Accuracy	Precision	Recall	F1 score	mIoU	Time per image
HD-Yolo	0.8263	0.7110	0.8308	0.6743	0.7409	0.8423	0.0108
HD-Staining	0.7724	0.6553	0.7943	0.6310	0.6994	0.8390	0.1503
Mask R-CNN	0.8000	0.6864	0.8215	0.6454	0.7198	0.8419	0.1704
HoVer-Net	0.7851	0.5867	0.6599	0.3953	0.5808	0.7717	1.0400

HD-Staining was processed with the repository and hyperparameters from the original paper. HoVer-Net was fine-tuned with the pretrained model from the PanNuke data set[38] released by the author. Mask R-CNN was upgraded from the official TorchVision Mask R-CNN by replacing the ResNet-50 FPN backbone with an EfficientNet-B3 feature pyramid network backbone for greater accuracy and faster speed. HD-Yolo achieved the fastest inference with the best performance.

HD-Yolo, Histology-based Detection using Yolo; mIoU, median intersection over union.

Table 2

Nucleus detection performance on the breast cancer multirater testing data set

Network	Backbone	Coverage	Accuracy	Precision	Recall	F1 score	Time per image
HD-Yolo	CSPDarkNet-l-Bi-FPN	0.8950	0.8042	0.9069	0.7992	0.8493	0.0089
Yolov7	E-ELAN	0.8723	0.8013	0.9065	0.7945	0.8468	0.0090
Yolov6	Efficient-Rep-PAN	0.8798	0.8133	0.9101	0.7963	0.8494	0.0088
Scaled-Yolov4	CSPDarkNet-PANet	0.8603	0.7977	0.8764	0.7810	0.8249	0.0123
FCOS	ResNet101-FPN	0.7876	0.7574	0.8848	0.7556	0.8148	0.0381
Deformable-DETR	ResNet50	0.7889	0.7620	0.8877	0.7588	0.8180	0.1420
EfficientDet	EfficientNetB3-Bi-FPN	0.8726	0.7950	0.8930	0.7936	0.8403	0.0951
Mask R-CNN	ResNet101-FPN	0.8631	0.7881	0.8749	0.7828	0.8252	0.1462
Cascade R-CNN	ResNeXt101_32x4d-FPN	0.8230	0.7915	0.9103	0.7902	0.8460	0.1720

HD-Yolo model is competitive with other advanced object detection algorithms, achieving blazing-fast inference speed. HD-Yolo, Histology-based Detection using Yolo.

Table 3

Performance of cross-validation study and final ensemble model on the breast cancer multirater inferred P-truth testing data set

	Coverage	Accuracy	Precision	Recall	F1 score	MCC	mIoU	Time per image
Model 1	0.8914	0.8017	0.9063	0.7960	0.8474	0.7246	0.7550	0.0042
Model 2	0.8891	0.7958	0.8989	0.7840	0.8373	0.7156	0.7545	0.0040
Model 3	0.8965	0.8092	0.9030	0.8026	0.8498	0.7335	0.8019	0.0041
Model 4	0.8995	0.8101	0.9012	0.8023	0.8478	0.7339	0.7725	0.0043
Model 5	0.8987	0.8050	0.8910	0.8050	0.8446	0.7288	0.7946	0.0041
Mean	0.8950	0.8044	0.9001	0.7980	0.8454	0.7273	0.7757	0.0041
Ensembled	0.9364	0.8655	0.9238	0.8655	0.8937	0.8044	0.8191	0.0197

Variance between models trained on different folds is rarely small. Ensemble modeling further improved performance by ~5% at the expense of computational speed. MCC, Matthews correlation coefficient; mIoU, median intersection over union.

Table 4

Inference speed comparison between different models on The Cancer Genome Atlas—Breast Cancer data set

Quantiles	Minimum	Median	95% quantile	Maximum
No. patches (512 × 512)	803	15,891	22,552	51,969
ROI diameter (×40, pixel)	16.38k	72.85k	86.78k	131.72k
HD-Yolo ensemble (h:min:s)	00:00:43	00:09:23	00:15:31	0:26:41
HD-Yolo single (h:min:s)	00:00:19	00:03:39	00:06:58	0:12:00
HD-Staining (h:min:s)	00:07:33	01:43:08	02:24:18	5:37:22
HoVer-Net (h:min:s)	00:15:50	04:46:38	11:51:22	26:38:12
TME feature extraction (h:min:s)	00:00:22	00:03:12	00:04:47	00:06:52
Export tables and figures (h:min:s)	00:01:09	00:29:15	00:32:56	00:45:58

HD-Yolo single model is about 25 to 30 times faster than HD-Staining and 50 to 130 times faster than HoVer-Net. The HD-Yolo ensemble model took double the time of its single counterpart with slightly better accuracy.

ROI, region of interest; TME, tumor microenvironment.