Annotation of Chorionic Villi in Products of Conception Using Combined Manual and Automated Techniques for Choriocarcinoma Research

Faraaz Rattani, Ashwin Menachery, Joshua J. Levy PhD
Emerging Diagnostic and Investigative Technologies, Department of Pathology, Dartmouth Hitchcock Medical Center

Dartmouth Health

ABSTRACT

•Choriocarcinoma– A rare and aggressive form of placenta cancer that can arise during or after pregnancy
•Traditionally, diagnosing choriocarcinoma relies on expert pathologists examining placental tissue for abnormal cells.
•The annotation process involved manually identifying chorionic villi within selected ROIs, using the "Segment Anything Model" (SAM) to assist with marking them.

INTRODUCTION

Choriocarcinoma Characteristics and Challenges

- Choriocarcinoma is a rare, aggressive placental cancer that can occur during or after pregnancy, making early detection difficult
- Symptoms often resemble common pregnancy issues, such as abnormal bleeding, leading to potential misdiagnosis
- Accurate diagnosis requires microscopic examination of placental tissues to detect choriocarcinoma cells, which can be challenging and time-consuming for pathologists

Traditional Diagnosis Methods

- Diagnosis has traditionally relied on expert pathologists' histopathological examination of placental tissue to detect abnormal cells
- These cells can be hidden within chorionic villi, necessitating careful and detailed analysis
- Even experienced pathologists may struggle due to the rarity of choriocarcinoma and limited case exposure, highlighting the need for advanced diagnostic tools

Research Project Focus

- The research aims to address diagnostic challenges by annotating high-resolution whole-slide images of products of conception (POC), identifying and marking each villus.
- An advanced segmentation tool, "Segment Anything," was used to enhance annotation accuracy.
- The project seeks to create a comprehensive dataset to support future AI-driven diagnostic tools, improving early detection and treatment of choriocarcinoma.

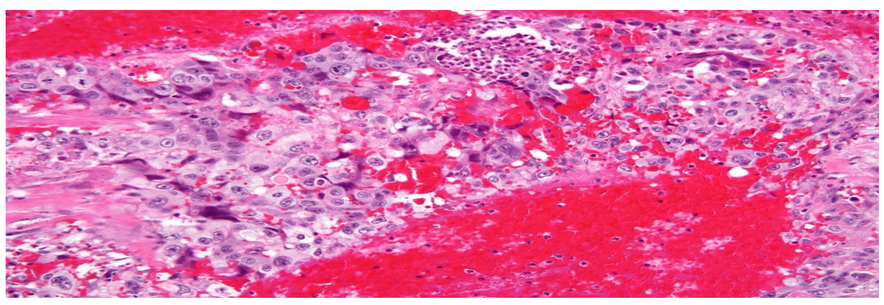

Figure 1: Image displaying choriocarcinoma

METHODS

- **Goal:** Investigate effectiveness of algorithmic methods in identifying lymphocytes from H&E that were tagged through imprecisely registered IF
- **Data Collection:**
  - High-resolution whole slide images (WSIs) of products of conception (POC) tissue samples were provided
  - The samples were sourced from Dartmouth Hitchcock Medical Center
  - The images included diverse villous structures for comprehensive analysis
- **Image Preparation:**
  - Whole slide images were examined to identify regions of interest (ROIs) containing significant villous structures
  - Each ROI was defined as a square with dimensions in multiples of 1024 pixels to ensure compatibility with future analysis tools
  - This approach facilitated detailed annotations and maintained uniformity across different samples
- **Annotation Process:**
- The annotation process involved manually identifying chorionic villi within selected regions of interest (ROIs).
- The "Segment Anything Model" (SAM) assisted in marking villi
- A standardized labeling system was used across all annotated images to support future analysis and machine learning model training.
- The annotation approach aimed to build a dataset that accurately reflects the complexity of chorionic villi in POC samples.
- Accurate identification focused on regions with key characteristics of placental villi, including syncytiotrophoblast cells, stromal cells, or evidence of capillaries.

RESULTS

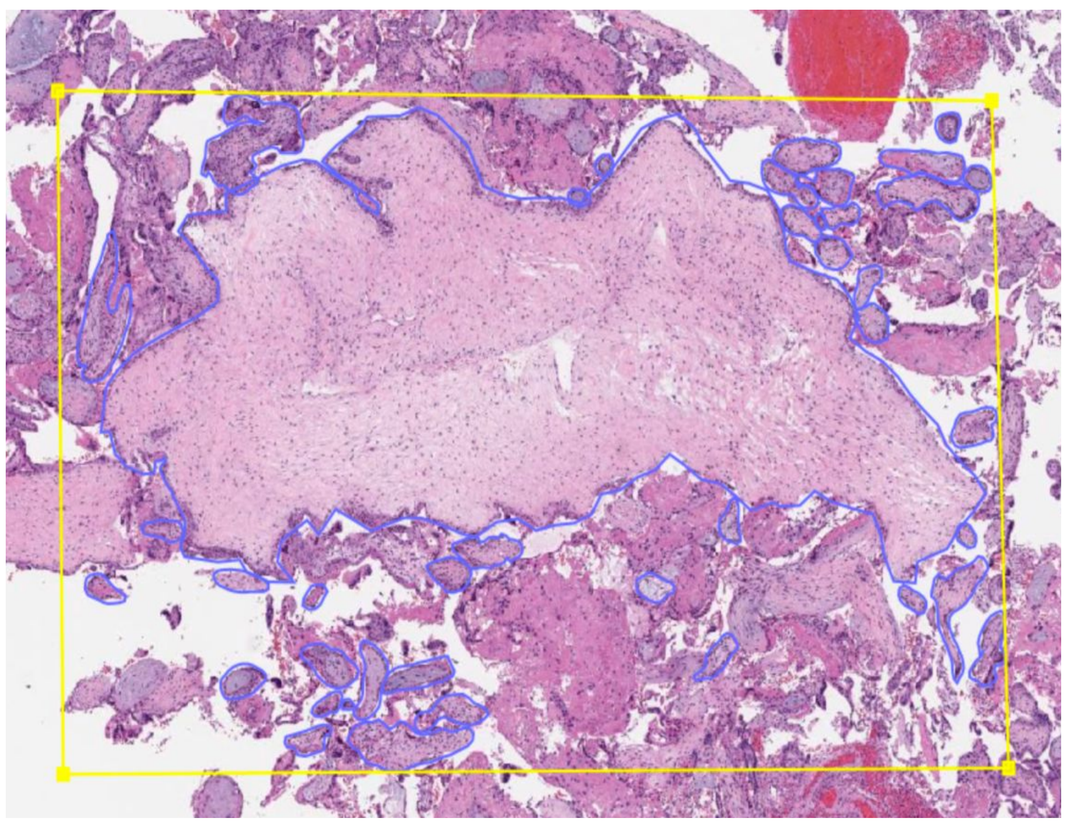

Figure 2: Annotated ROI of a whole slide image with villi marked in blue

- Annotated approximately 4,000 chorionic villi from high-resolution whole-slide images of POC tissue samples.
- Encountered challenges due to irregular and varied shapes of the villi.
- Villi shapes and sizes were random, with no discernible patterns, making consistent identification difficult.
- Used the "Segment Anything Model" (SAM) for annotation, which generally helped but struggled with unpredictable villi shapes and indistinct boundaries.
- Manual annotation was necessary when SAM failed to accurately mark villi, ensuring dataset accuracy and reliability.
- Despite challenges, the combined SAM and manual annotation approach produced an accurate and reliable annotated dataset.

CONCLUSION

- **Potential for Clinical Impact:**
  - Enhances the accuracy of diagnosing placental disorders, including choriocarcinoma, by aiding pathologists in identifying abnormal structures
  - Supports the early detection of placental conditions, like choriocarcinoma, leading to better patient outcomes
  - Helps reduce the manual workload for pathologists, enabling faster and more consistent analysis of placental tissues
- **Limitations:**
  - Annotating the whole slide images take significant amounts of time
  - Room for inaccuracies due to misclassification of villi and human error
- **Future Directions:**
  - Future work involves using annotated data in models for further analysis. First, use a U-net model for image segmentation to accurately distinguish chorionic villi from the background and other tissues. U-net model is trained on annotated images to recognize and outline precise boundaries of the villi. Next, apply a Convolutional Neural Network (CNN) for feature analysis. CNN examines segmented images to identify features indicating choriocarcinoma. CNN extracts and learns from complex patterns in the villi to detect potential malignancy.
- **Data and Code Availability:**
  - Data/code available on reasonable request, privacy/ethical restrictions.

Acknowledgements: EDIT, DPLM, Pathology Shared Resource, DCC @ DHMC.

edit.