

Map Understanding Model: Generating GeoSpatial Linked Data from Map Images

2023 AAG Annual Meeting

Jina Kim, Yao-Yi Chiang

Knowledge Computing Lab

Department of Computer Science and Engineering, University of Minnesota

Motivation



Post Route Map of the State of Minnesota Showing Post Offices ... October 1, 1944.

<https://www.davidrumsey.com/luna/servlet/detail/RUMSEY~8~1~321690~9009088~8?cic=RUMSEY%7E8%7E1#>

- Many scanned maps are available
- The lack of **metadata**
 - Difficult to search and find the right maps
- **The content of maps** are meaningful resource

Motivation



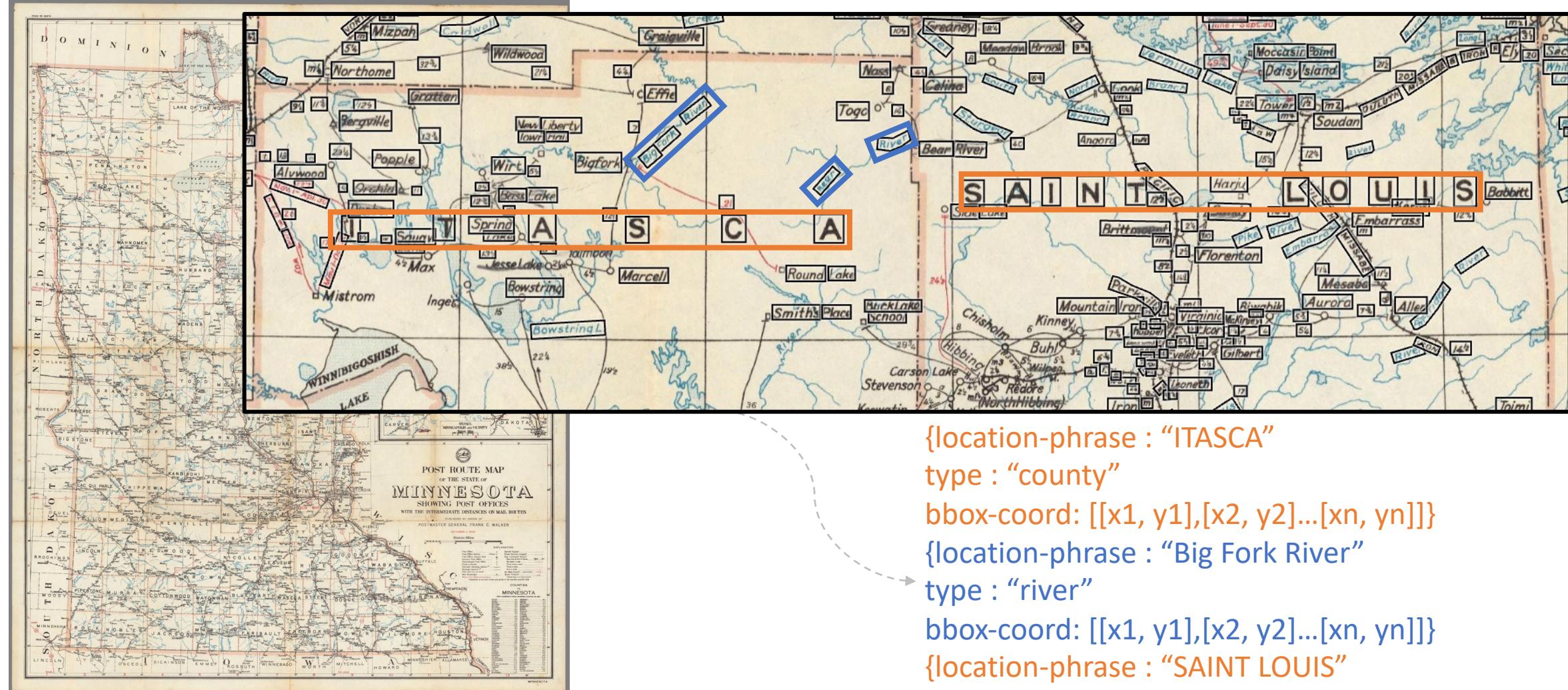
Post Route Map of the State of Minnesota Showing Post Offices ... October 1, 1944.

<https://www.davidrumsey.com/luna/servlet/detail/RUMSEY~8~1~321690~9009088~8?cic=RUMSEY%7E8%7E1#>

- Many scanned maps are available
- The lack of **metadata**
 - Difficult to search and find the right maps
- **The content of maps** are meaningful resource

RQ. How can we automatically extract structured textual information from scanned maps?

Expected Results



{location-phrase : "ITASCA"
 type : "county"
 bbox-coord: [[x1, y1],[x2, y2]...[xn, yn]]}
 {location-phrase : "Big Fork River"
 type : "river"
 bbox-coord: [[x1, y1],[x2, y2]...[xn, yn]]}
 {location-phrase : "SAINT LOUIS"
 type : "county"
 bbox-coord: [[x1, y1],[x2, y2]...[xn, yn]]}

Previous Work on Document Layout Analysis

- Document understanding models **jointly understand the document layout and contents** [LayoutLMv1,2,3, BROS, SPADE]

SPORTS MARKETING ENTERPRISES
DOCUMENT CLEARANCE SHEET

Date Routed: January 11, 1994 Contract No. 4011 00 00

Contract Subject: Joe's Place Exhibits

Company SPEVCO, INC. Brand(s) Camel/Winston

Total Contract Cost \$1,340,000.00 Current Year Cost 1994-1995

Brief Description 2 Joe's Place Exhibits for use at Winston Cup, Winston Drag and Camel Super Bike Events.

G.R. Code: Program Budget Code

Originator Michael Wright

Manager John Powell B. J. Powell 1-11-94

REVIEW ROUTING

Insurance _____

Law _____

FS - Marketing _____

REVISIONS TO SHELL (Other than Term, Compensation or Job)

PAGE(S) SECTION(S)

APPROVAL ROUTING * Sr. Manager (B. J. Powell)
* Director - (G. L. Littell)

** Sr. VP T. W. Robertson

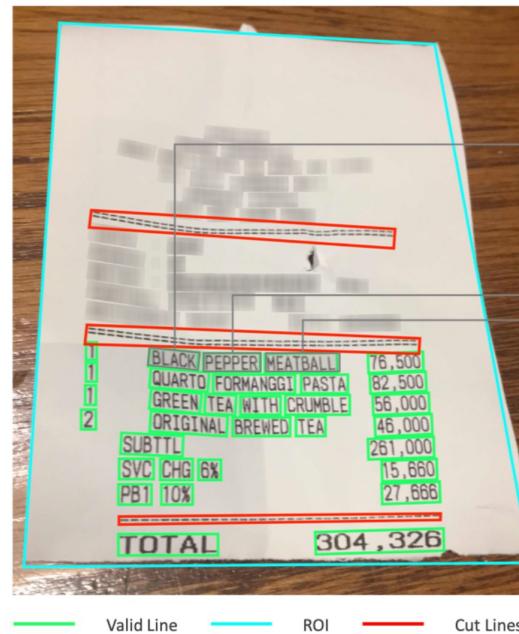
Return To: MARY SEAGRAVES Ext. 1485 SME 13 Plaza

* UP TO AND INCLUDING \$25,000
**OVER \$25,000

Revised 10/26/92

Example of FUNSD dataset

(<https://guillaumejaume.github.io/FUNSD/>)



Image

Example of CORD dataset

(<https://github.com/clovaai/cord>)

```
...
"words": [
  {
    "quad": {"x2": 837, "y3": 2260, "x3": 829, "y4": 2238,
             "x1": 625, "y1": 2156, "x4": 617, "y2": 2177},
    "is_key": 0,
    "row_id": 1631316,
    "text": "BLACK"
  },
  {
    "quad": {"x2": 1105, "y3": 2261, "x3": 1105, "y4": 2261,
             "x1": 857, "y1": 2180, "x4": 857, "y2": 2180},
    "is_key": 0,
    "row_id": 1631316,
    "text": "PEPPER"
  },
  {
    "quad": {"x2": 1441, "y3": 2264, "x3": 1443, "y4": 2269,
             "x1": 1125, "y1": 2184, "x4": 1126, "y2": 2179},
    "is_key": 0,
    "row_id": 1631316,
    "text": "MEATBALL"
  }
],
"category": "menu_nm",
"group_id": 3
}
...
}
```

JSON

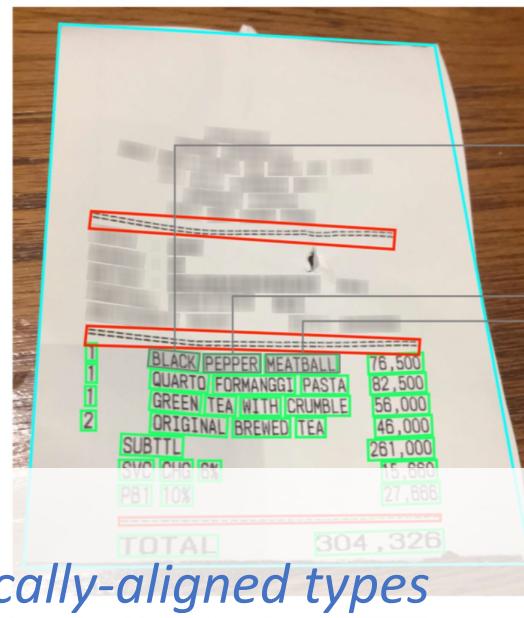
Previous Work on Document Layout Analysis

- Document understanding models **jointly understand the document layout and contents** [LayoutLMv1,2,3, BROS, SPADE]

A scanned document titled "SPORTS MARKETING ENTERPRISES DOCUMENT CLEARANCE SHEET". It contains several sections of text and checkboxes. A large black arrow points from the bottom right towards the "Vertical alignment types" section below.

Date Routed: January 11, 1994 Contract No. 4011 00 00
Contract Subject: Joe's Place Exhibits
Company SPEVCO, INC. Brand(s) Camel/Winston
Total Contract Cost \$1,340,000.00 Current Year Cost 1994-1995
Brief Description 2 Joe's Place Exhibits for use at Winston Cup, Winston Drag and Camel Super Bike Events.
G.R. Code: Program Budget Code
Originator Michael Wright
Manager John Powell B. J. Powell 1-11-94
REVIEW ROUTING
Insurance
Law
FS - Marketing
REVISIONS TO SHELL (Other than Term, Compensation or Job)
PAGE(S) SECTION(S)
APPROVAL ROUTING * Sr. Manager (B. J. Powell)
* Director - (G. L. Littell)
** Sr. VP T. W. Robertson
Return To: MARY SEAGRAVES Ext. 1485 SME 13 Plaza 38
* UP TO AND INCLUDING \$25,000
** OVER \$25,000
Revised 10/26/92

Example of FUNSD dataset
(<https://guillaumejaume.github.io/FUNSD/>)



Vertically-aligned types
How about complex documents such as historical maps?

Image

```
{  
  "words": [  
    {  
      "quad": {"x2": 837, "y3": 2260, "x3": 829, "y4": 2238,  
               "x1": 625, "y1": 2156, "x4": 617, "y2": 2177},  
      "is_key": 0,  
      "row_id": 1631316,  
      "text": "BLACK"  
    },  
    {  
      "quad": {"x2": 1105, "y3": 2261, "x3": 1105, "y4": 2261,  
               "x1": 857, "y1": 2180, "x4": 857, "y2": 2180},  
      "is_key": 0,  
      "row_id": 1631316,  
      "text": "PEPPER"  
    },  
    {  
      "quad": {"x2": 1441, "y3": 2264, "x3": 1443, "y4": 2269,  
               "x1": 1125, "y1": 2184, "x4": 1126, "y2": 2179},  
      "is_key": 0,  
      "row_id": 1631316,  
      "text": "MEATBALL"  
    },  
  ],  
  "category": "menu_nm",  
  "group_id": 3  
}
```

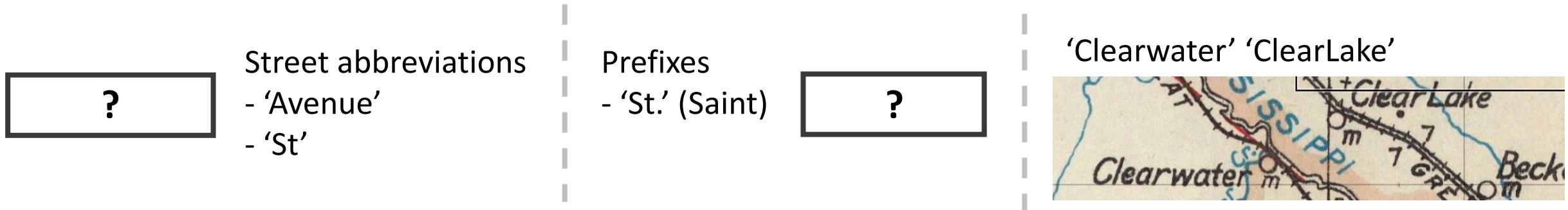
JSON

Example of CORD dataset
(<https://github.com/clovaai/cord>)

Relations of Text Labels in Map Images (1/3)

- Text relations
 - capture the **semantic meaning and word order** within the location phrase

Examples

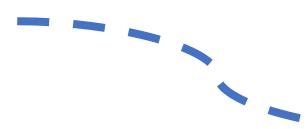


Relations of Text Labels in Map Images (2/3)

- Semantic type relations
 - capture the **relationship between semantic types** of multiple location phrases

Example

'beach' 'lake' and 'park'



Relations of Text Labels in Map Images (3/3)

- Visual relations
 - capture the **visual font style** (e.g., height, font family) of the location phrase
- Spatial relations
 - capture the **relative positions and orientation** of words within the location phrase

Example

'RED LAKE' (county)

'Red Lake River' (river)

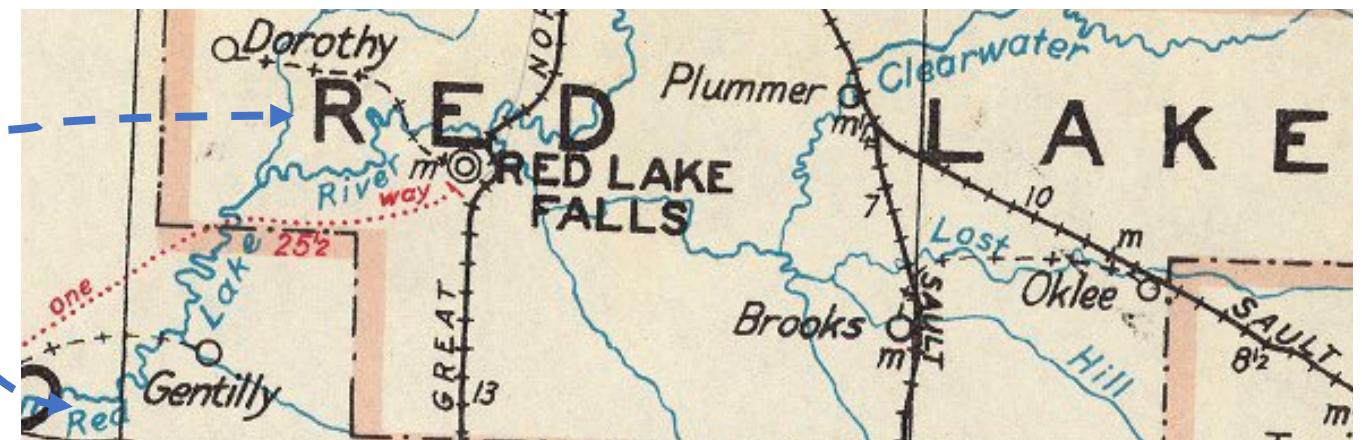


Figure from Post Route Map of the State of Minnesota Showing Post Offices ... October 1, 1944.

<https://www.davidrumsey.com/luna/servlet/detail/RUMSEY~8~1~321690~90090888?cic=RUMSEY%7E8%7E1#>

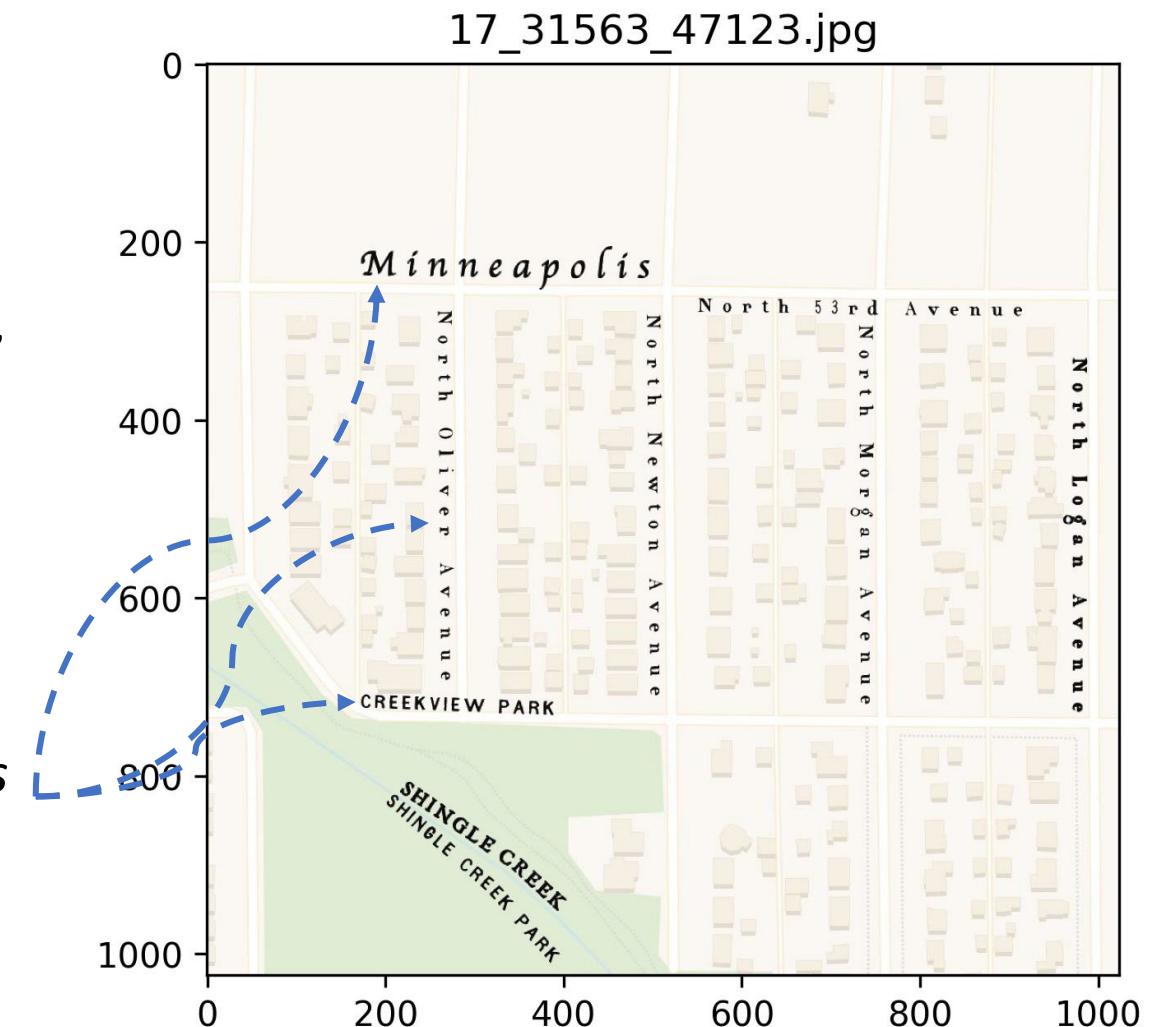
Verifying Document Understanding Model's Ability

- Hypothesis
 - Document understanding model can capture **some relations of text labels on synthetic maps**
- Experiment settings
 - Task definition - predict **semantic type of each word**
 - Model - LayoutLMv2 (Xu et al., 2021)
 - Dataset – SynthMap
 - Semantic types - buildings, roads, railways, water, places (e.g., city, town)

SynthMap Dataset

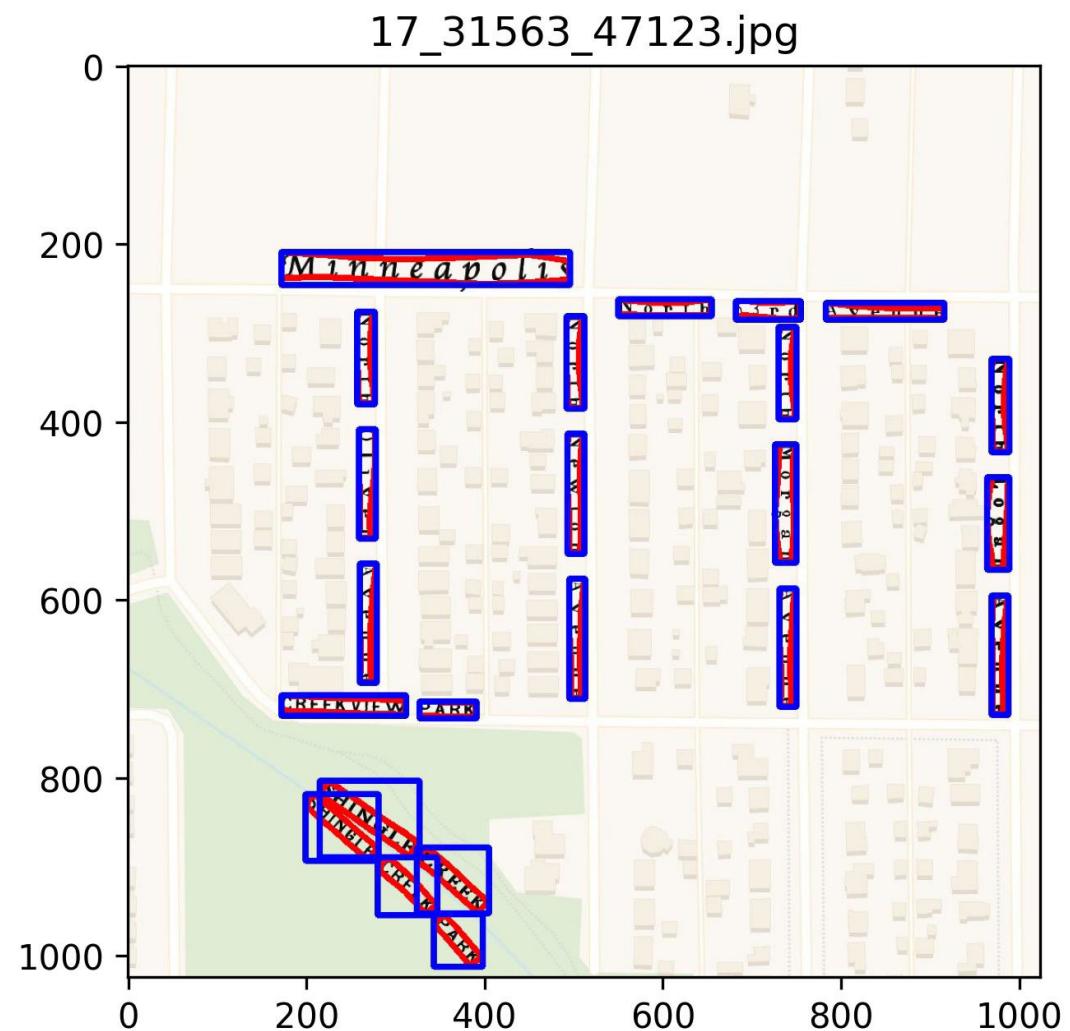
- Map image (JPG)
- Text label information (JSON)
 - fields: name, word, word_contour, osm_id, geo_feature, feature_type

Font style and font size differs from semantic types



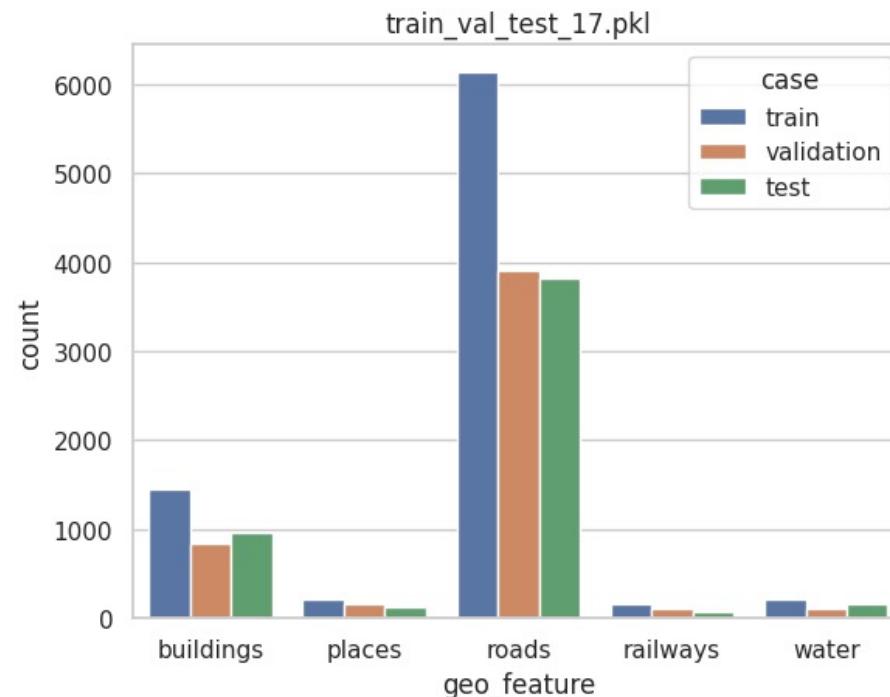
SynthMap Dataset

- Map image (JPG)
- Text label information (JSON)
 - fields: name, word, **word_contour**, osm_id, geo_feature, feature_type, **word_bounding_rect** (for LayoutLMv2)



Experiment Results

- # of images - train 404 / val 253 / test 253
- # of words - train 8,167 / val 5,111 / test 5,131



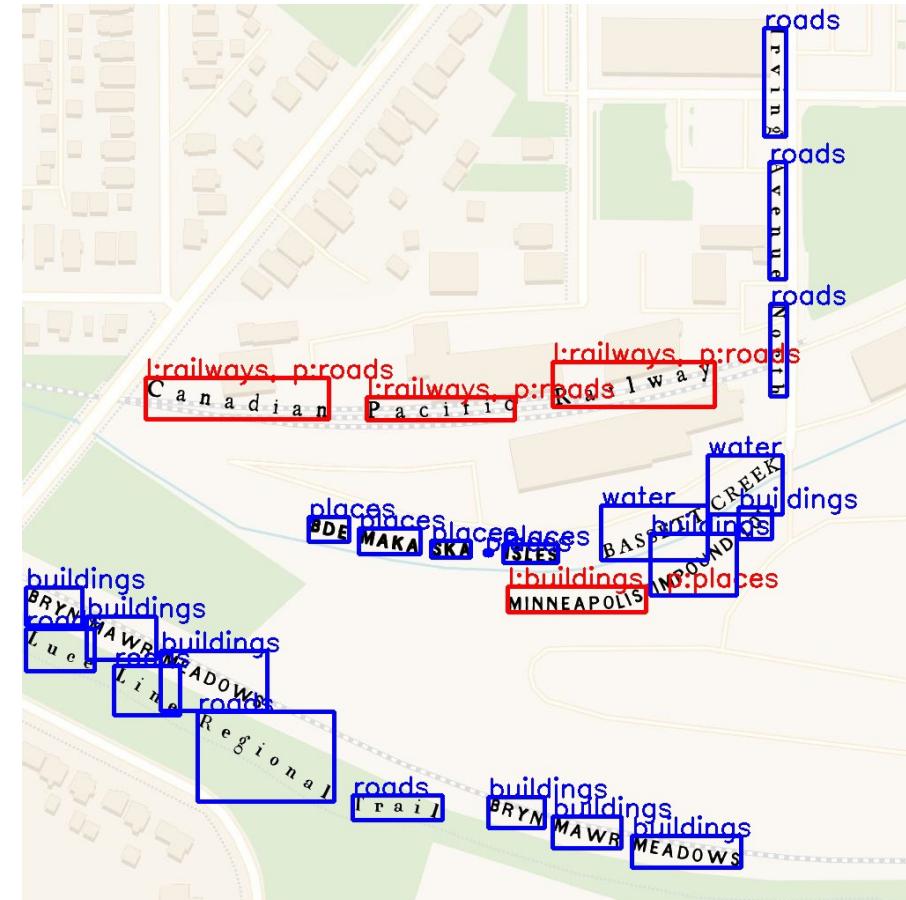
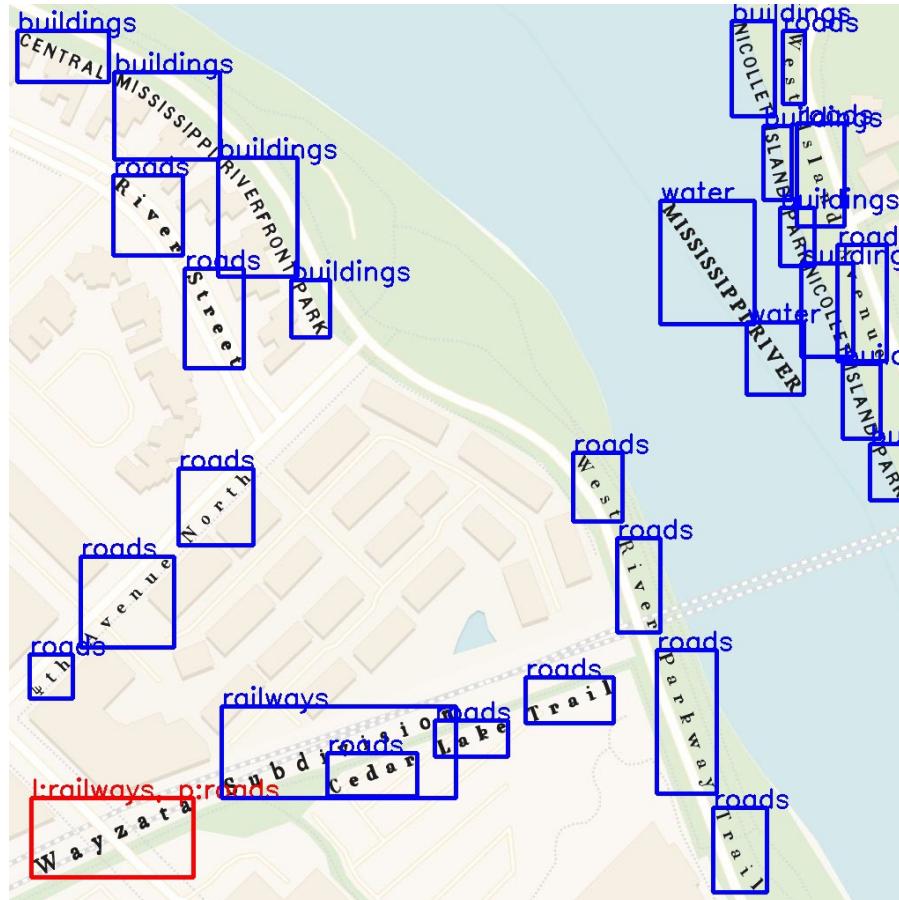
A descriptive analysis on datasets

	Type	F1 ↑	Weighted F1 ↑
LayoutLMv2	Railways	93.33	
	Water	96.61	
	Roads	98.09	97.70
	Places	95.92	
	Buildings	98.19	

Model performance on test set

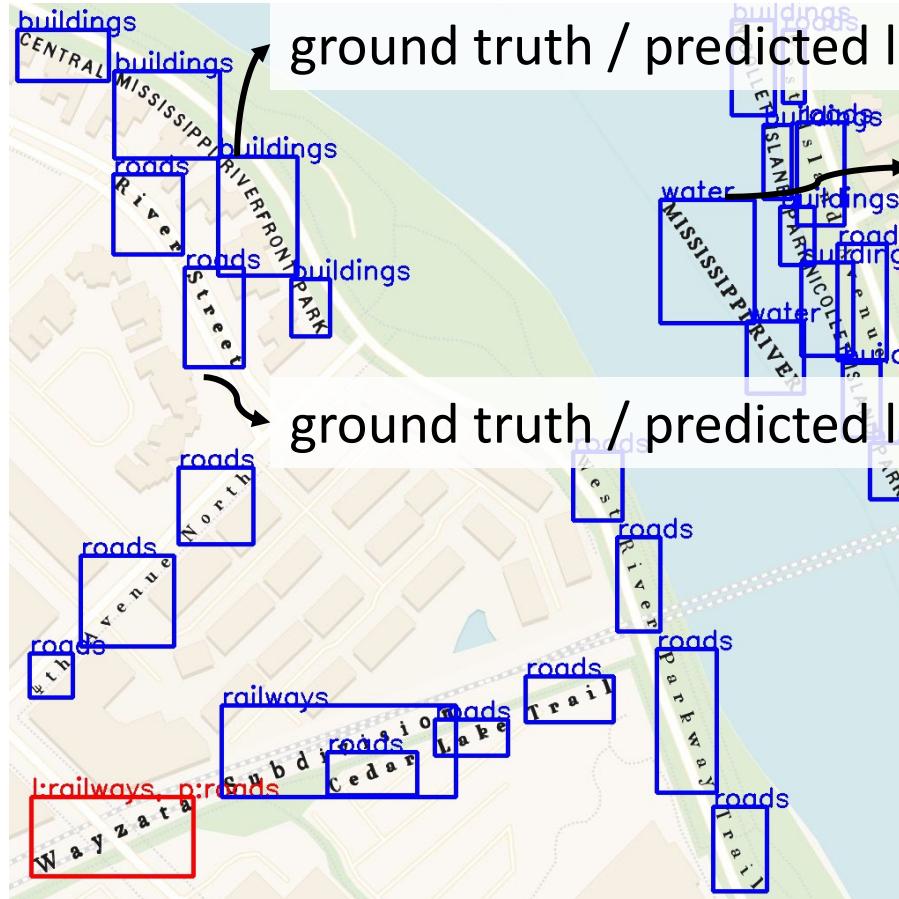
Prediction Visualization

- Blue - correctly predicted type | Red - incorrectly predicted type



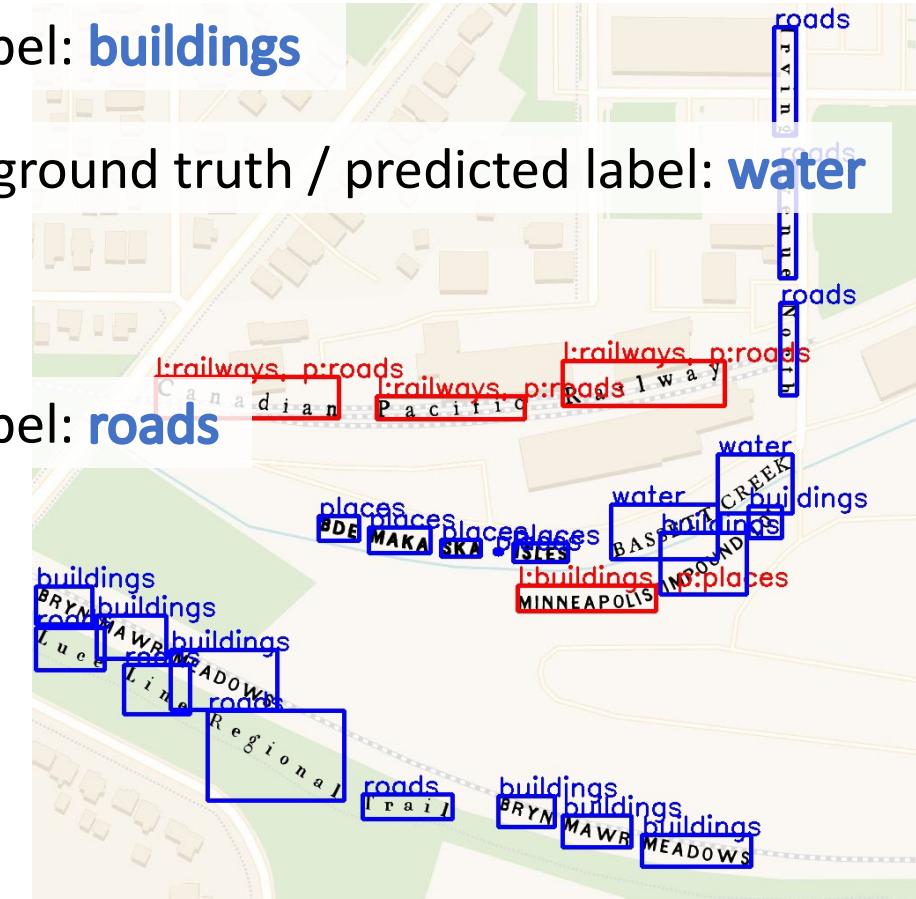
Prediction Visualization

- Blue - correctly predicted type | Red - incorrectly predicted type



ground truth / predicted label: **buildings**

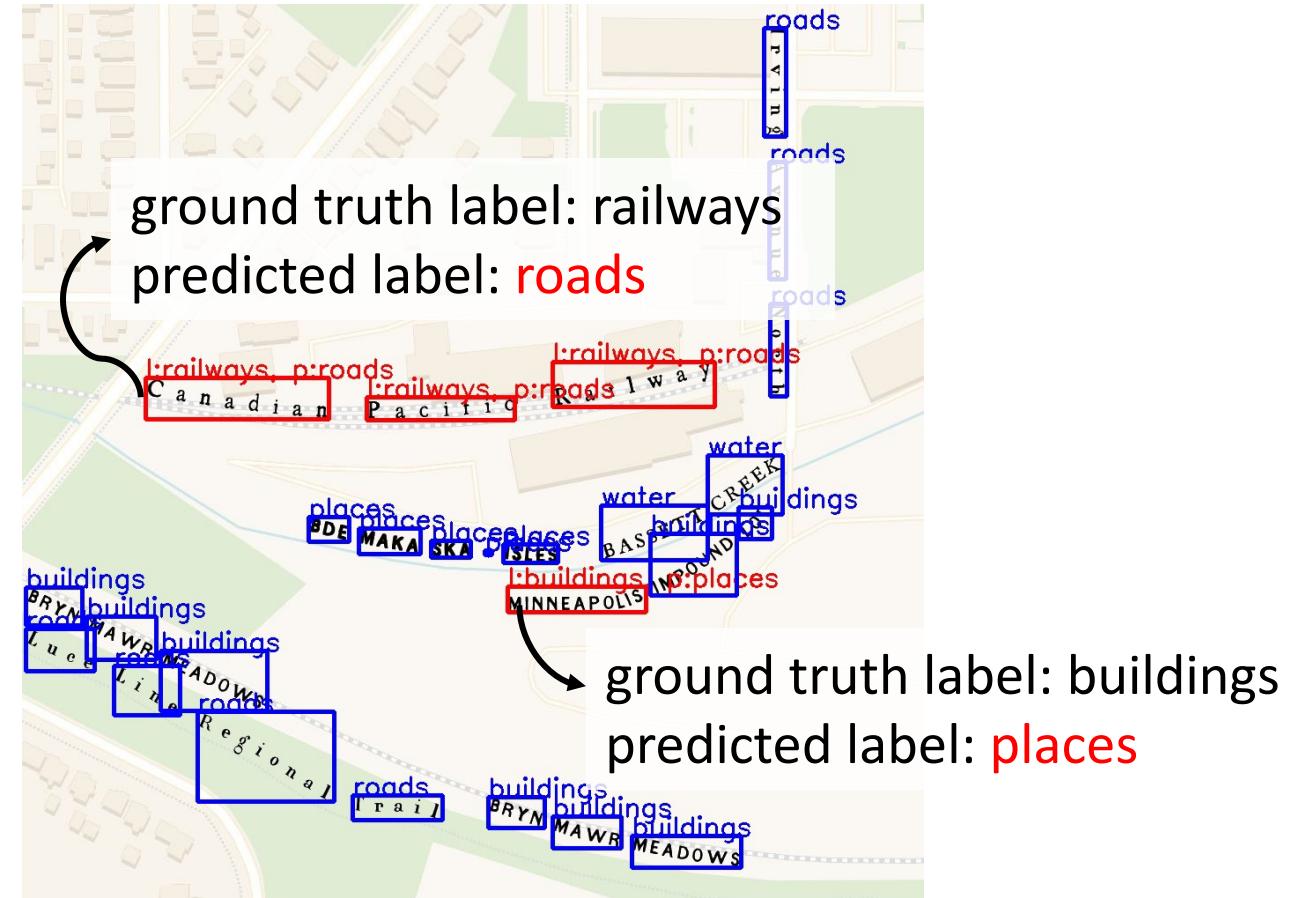
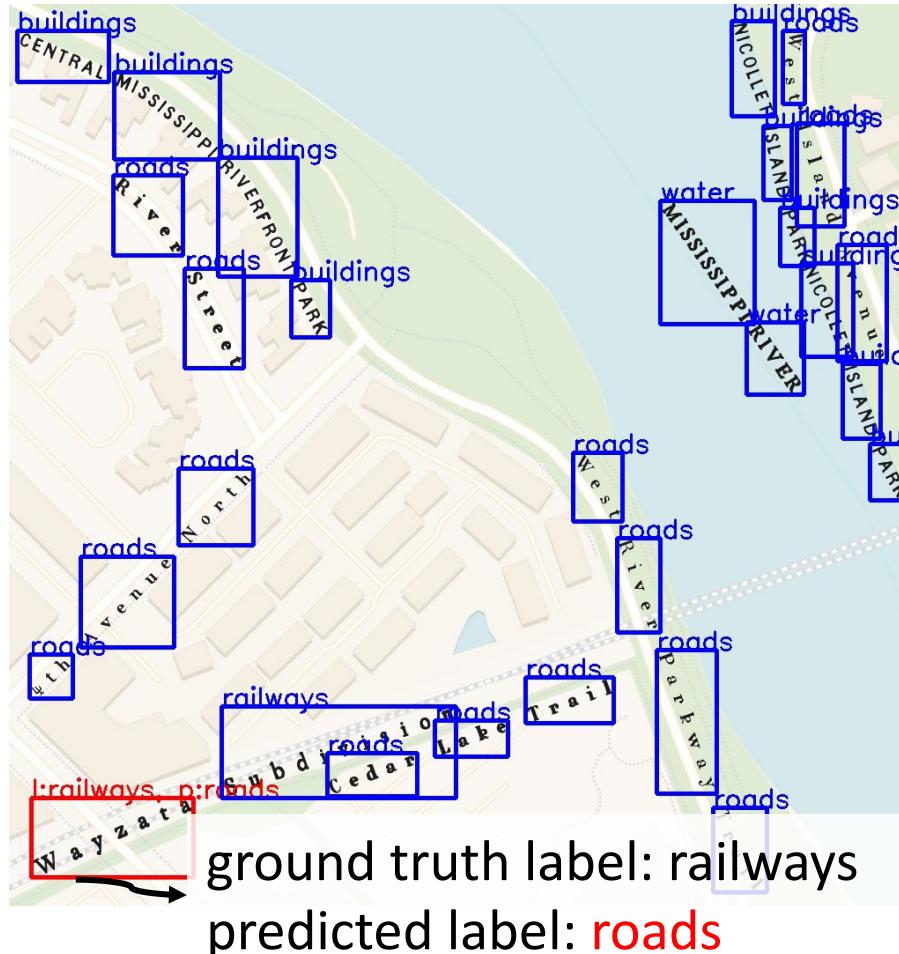
ground truth / predicted label: **roads**



ground truth / predicted label: **water**

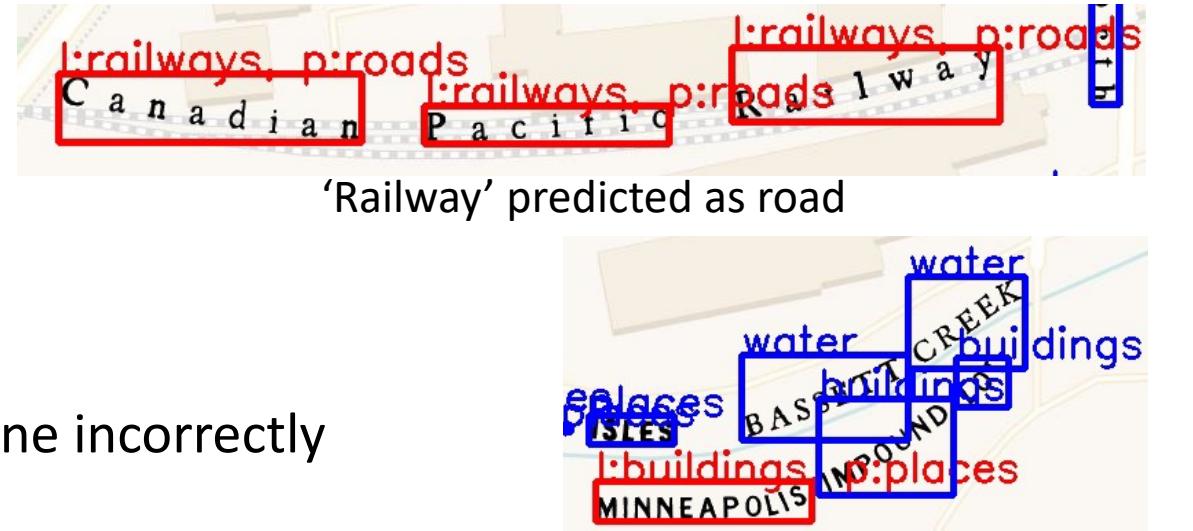
Prediction Visualization

- Blue - correctly predicted type | Red - incorrectly predicted type

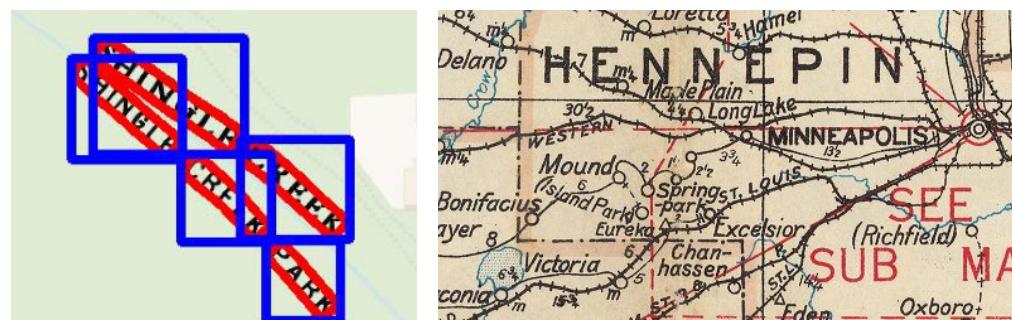


Result Analyses & Remaining Challenges

- Result analyses on four relations
 - Text relations (some)
 - Semantic type relations (some)
 - Visual relations (very well)
 - Spatial relations (some)
 - 11 location phrases have more than one incorrectly predicted words
- Remaining challenges
 - Larger background area in bounding box with curved text labels
 - Hard to capture relations with complex backgrounds of historical maps



'MINNEAPOLIS' predicted as place



A comparison of synthetic map and real scanned map

Takeaways

- **Text, semantic type, visual, and spatial relations** are important cues for map understanding
- Previous model **can capture visual relations** of text labels on synthetic maps very well
- To address complex backgrounds of real scanned maps, **new technology among four relations are required**