

Towards automatic parsing of arbitrary form images

The Label-Value pairing problem

Brian Davis

Work done in collaboration with

Bill Barrett, Bryan Morse, Chris Tensmeyer, Scott Cohen, Brian Price

BYU

Adobe Research



BRIGHAM YOUNG UNIVERSITY

Family History

TECHNOLOGY WORKSHOP 2019



Outline

- Problem definition
- Previous work
- Dataset
- Elements of a solution
- Subproblem: Label-Value pairing
- Method
 - Detection
 - Pairing
 - Optimization
- Results
- Conclusion

The problem

- Genealogical data is often recorded on forms
 - Birth, death, marriage records, etc.
- We'd like to automatically extract this information

The problem

- Some collections are not of uniform form types
 - Most have variations of the same form
 - Some are fairly random assortments for forms
- We want to be able to automatically extract information from a form image that we have never seen before
 - **Template-free**
 - **Noisy, historical forms**

Elements of a solution

1. Detect (locate) pre-printed text and input writing on the page
2. Transcribe (read) the pre-printed text and input writing
3. Identify the meaning of the input writing (using the pre-printed text and layout)
4. Format the transcription of the handwriting to match a data schema

The data we care about

CHECK TYPE REQUIRED
(See Instructions on back)

APPLICATION FOR HEADSTONE OR MARKER
(If from state and find return in duplicate)

UPRIGHT MARBLE HEADSTONE
 UPRIGHT GRANITE HEADSTONE
 FLAT MARBLE MARKER
 FLAT GRANITE MARKER
 BRIDGE MARKER (NOTE RESTRICTIONS)

ENLISTMENT DATE: April 28, 1917
 DISCHARGE DATE: Feb. 3, 1919

SERIAL NO.: 295, 607
 RANK: none
 COMPANY: Illinois Sargeant 6--128th
 U. S. REGIMENT, STATE ORGANIZATION AND DIVISION: Illinois Sargeant 6--128th

NAME (Last, First, Middle Initial): Enalin, Irving David
 DATE OF BIRTH (Month, Day, Year): July 25, 1892
 DATE OF DEATH (Month, Day, Year): Oct. 29, 1943
 NAME OF CEMETERY: Mount Hope

NEAREST FREIGHT STATION (City and State): Belleveille, Illinois
 POST OFFICE ADDRESS OF CONSIGNEE: 2008 - Washington St. Belleville, Illinois

SIGNATURE OF CONSIGNEE: D. Enman
 DO NOT WRITE HERE
 FOR VERIFICATION: FER 28 1947
 COMMISSIONER: WELMSFORD, MASS. APR 9 1947
 ADDRESS (Street, City, State): 915 Pennsylvania Ave; E. St. Louis, IL

QDMS FORM 623
 REV 4 NOT 4

To the Honorable Collector of Customs,
Baltimore, Md.

Sir: I respectfully apply for a certificate of Departure as provided for under the act of September 13, 1899, and the treaty between the United States of America and the Empire of China, signed March 17, 1894, as I am a Chinese laborer residing in the United States, and have a lawful wife, child or parent therein, or property therein of the value of \$1,000, or debts of like amount due me and pending settlement, as hereinafter set forth.

My wife if any and hereinafter named, was married to me at least one year, and since then we have continuously cohabited as man and wife. My property, if any, and hereinafter set forth, is bona fide and not of credit and sale only.

My debts due me, if any and hereinafter set forth, are unascertained notes or other similar acknowledgments of unsecured liability.

At least a month from date hereof I intend to depart to China via the port with the intention of returning to my port within one year from such date of departure.



MY DESCRIPTION
 Name: Yung Yung
 Local Residence: 217 N. 4th St. Baltimore, Md.
 Age: 45 years
 Occupation: laborer
 Height: 5 feet
 Weight as present: 140 lbs.
 Color of Eyes: Black
 Complexion: Dark
 Physical Marks: Very

No. Certificate of Residence: under left eye

Family Property or Debts: 22671

Sam Chung 214 N. 10th

Miss Kim 600.00

Lee Gove 927 Race St. Philadelphia Pa.

Miss Kim 500.00

楊成基

Respectfully yours,

Yung Yung
Wife: Mabel

Subscribed and sworn to before me this

Y. J. McQuinn
Deputy Collector of Customs
Baltimore, Md.

I HEREBY CERTIFY, That after a thorough examination I am satisfied that the foregoing applicant is registered, ~~and that his height, weight and description of physical marks are accurately given.~~ That the photograph hereto attached is a correct and faithful likeness of him, and that his height, weight and description of physical marks are accurately given.

Witness my hand and official seal this

Y. J. McQuinn
Deputy Collector of Customs
Baltimore, Md.

Page No. 3
SCHEDULE 1 - Free Inhabitants in Baltimore **in the County of** Stearns **State**
of Alabama **enumerated by me, on the** 31 **day of** June **1880.** W. H. K. K. K. **Asst. Marshal**
Post Office Washington

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | | 9 | 10 | 11 | 12 | 13 | 14 |
|---|---|---|---|---|---|---|----------------------|--------------------------|----|----|----|----|----|----|
| | | | | | | | Value of Real Estate | Value of Personal Estate | | | | | | |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |
| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 |

N 25 - 146
 Family name: WIEGAND Given name: ADAM
 Accompanied by: (COVER)
 Age: 42 Sex: M. S. Occupation: LAND MAN Nationality: ELLINGSHOUSEN
 Last permanent residence (town country, etc.): BALTIMORE, MD. Destination: 1847
 Name of vessel: LOUISE SOPHIE Date: MAY 24
 DEPARTMENT OF LABOR, Immigration and Naturalization Service, Form 548-C

This is form of usually signs the seal. Evidence that is verified and agree to accept the above of this seal ports covering on a Act in confirmation of seal. That my name or name of party, due the seal Evidence under the Agreement, shall constitute a alignment of problem shall be made until the duly authorized Agent shall certify that it does to balance on paid or satisfactory ground.

No Restitution Warranted, the said parties have affixed their names to this Agreement in presence

| NO. | NAMES | AGE | SEX | CLASS | Number of previous entries as Laborer or Immigrant | Rate of Wages | Height | Weight | Complexion | REMARKS | NO. |
|-----|-------------------|-----------|-----------|----------|--|---------------|--------|--------|------------|---------|-----|
| 1 | <u>Frank Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 2 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 3 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 4 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 5 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |

| | | | | | | | | | | | |
|----|------------------|-----------|-----------|----------|--|--|--|--|--|--|--|
| 24 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 25 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 26 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 27 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 28 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 29 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 30 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 31 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 32 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 33 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 34 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 35 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 36 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 37 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 38 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 39 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |
| 40 | <u>John Wood</u> | <u>23</u> | <u>M.</u> | <u>1</u> | | | | | | | |

Dataset

- From the National Archive, imaged by FamilySearch
- We restrict our problem to forms without tables or fill-in-the-blank paragraphs
 - Though we use the whole set to pre-train our detector
- Size
 - Train: 143 images (full 682)
 - Validation: 13 images (full 59)
 - Test: 13 images (full 62)

CHECK TYPE REQUIRED
(See instructions attached)
 TOPRIGHT GRANITE HEADSTONE
 TOPRIGHT GRANITE HEADSTONE
 FLAT MARBLE MARKER
 FLAT GRANITE MARKER
 OTHER MATERIALS (NOTE RESTRICTIONS)

APPLICATION FOR HEADSTONE OR MA
(Please make out and return in duplicate)

ENLISTMENT DATE: 26 May 1944
DISCHARGE DATE: 8 Feb. 1946

SERIAL NO.: 971 05 45

NAME (Last, First, Middle Initial): Ansley, Buford L.

DATE OF BIRTH (Month, Day, Year): July 24, 1925
DATE OF DEATH (Month, Day, Year): June 23, 1949

NAME OF CEMETERY: Rose Hill Cemetery

SIGNATURE OF CONSIGNEE: Robert F. Young

FOR VERIFICATION: JUN 15 1949
DISCHARGE: Columbus, Miss. 8 SEP 1949

ADDRESS: Carters Creek, P.O., Tennessee

DO NOT WRITE HERE

DOING FORM REV 1 NOV 48 623

IMPORTANT—Complete Reverse Side

Talk to me if you're interested in using it.

Subproblem: Label-Value pairing

Often determining the meaning of input writing is as simple as knowing it's label

BRONZE MARKER (NOTE RESTRICTIONS)

NAME (Last, First, Middle Initial)

Ensley, Buford L. *marker*

DATE OF BIRTH (Month, Day, Year)

July 24, 1925

DATE OF DEATH (Month, Day, Year)

June 28, 1949

NAME OF CEMETERY

Rose Hill Cemetery

SHIP TO (I CERTIFY THE APPLICANT FOR THIS STONE HAS MADE ARRANGEMENTS WITH ME TO TRANSPORT

Subproblem: Label-Value pairing

Often determining the meaning of input writing is as simple as knowing it's label

BRONZE MARKER (NOTE RESTRICTIONS)

NAME (Last, First, Middle Initial)

Ensley, Buford L. *marker*

DATE OF BIRTH (Month, Day, Year)

July 24, 1925

DATE OF DEATH (Month, Day, Year)

June 28, 1949

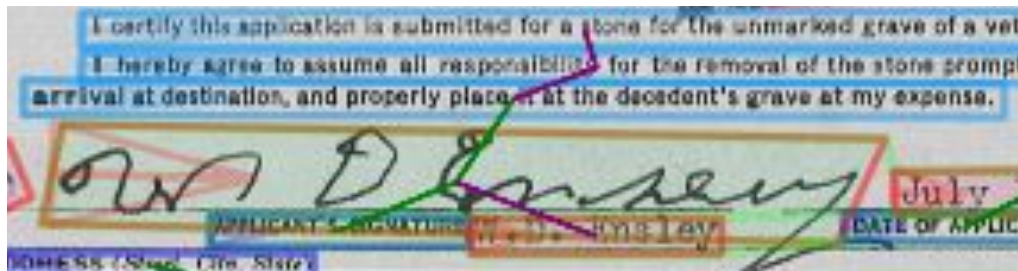
NAME OF CEMETERY

Rose Hill Cemetery

SHIP TO (I CERTIFY THE APPLICANT FOR THIS STONE HAS MADE ARRANGEMENTS WITH ME TO TRANSPORT

Subproblem: Label-Value pairing

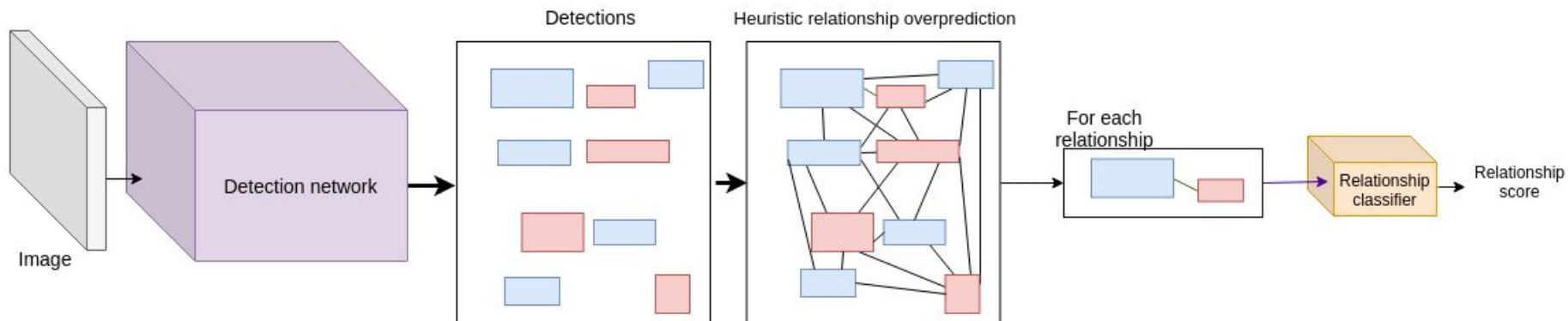
- First pre-printed text and input writing must be detected/located
- Then we need to decide which pre-printed text is related to which input writing
 - Not always one-to-one relationships!



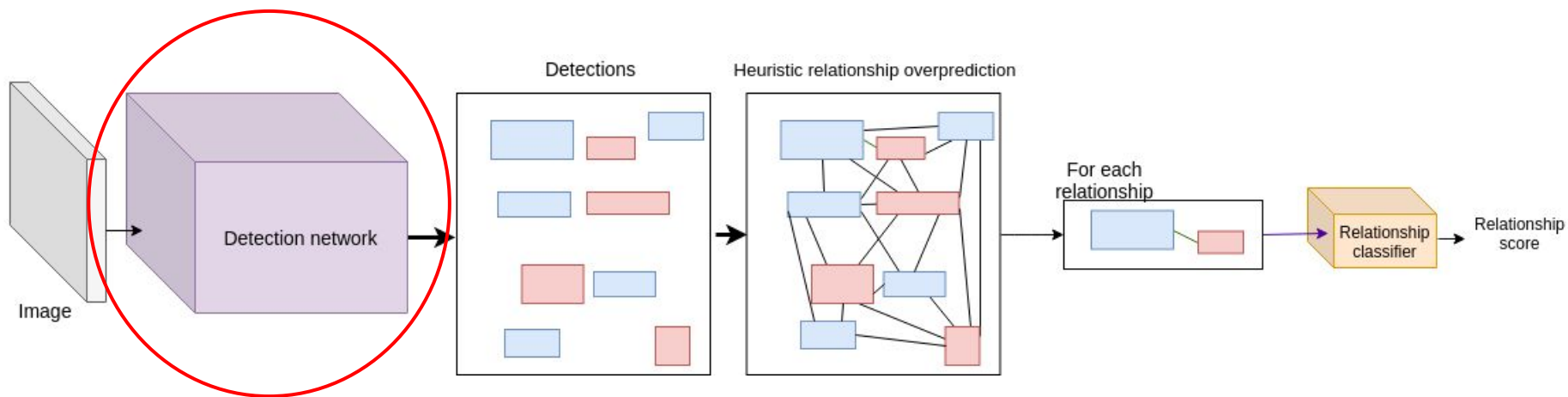
We do this without transcribing/reading the text and handwriting.

- Language agnostic

Method

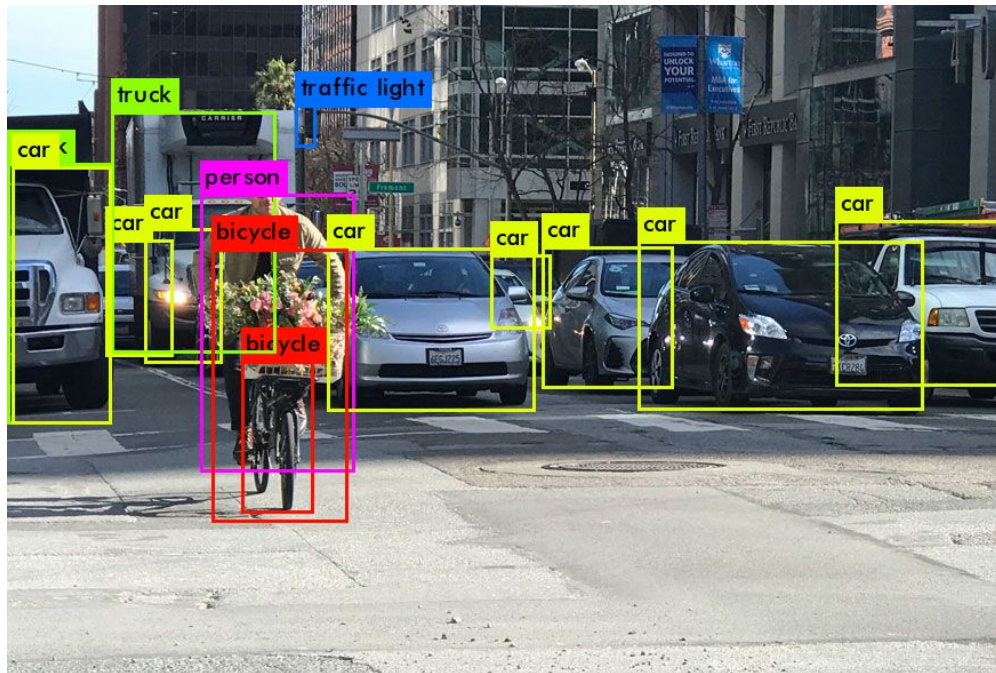


Method



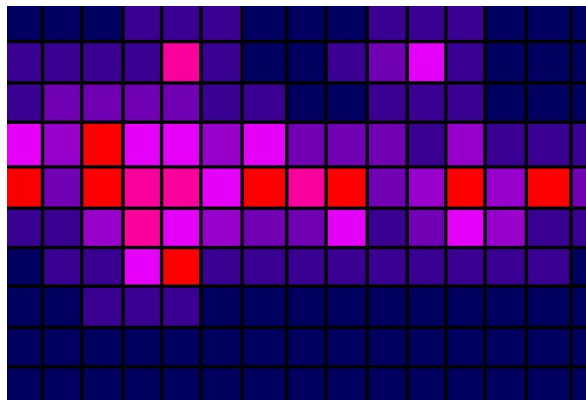
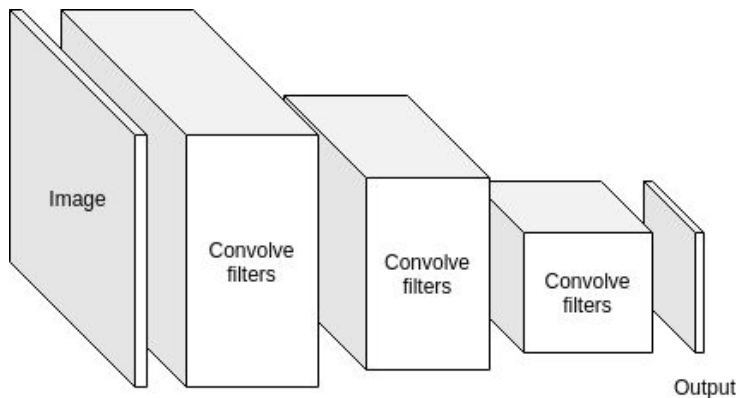
Detection

- We detect lines of text/writing using an object detection method.
- Specifically we model our detector after YOLOv2/v3.
 - Single fully convolutional network



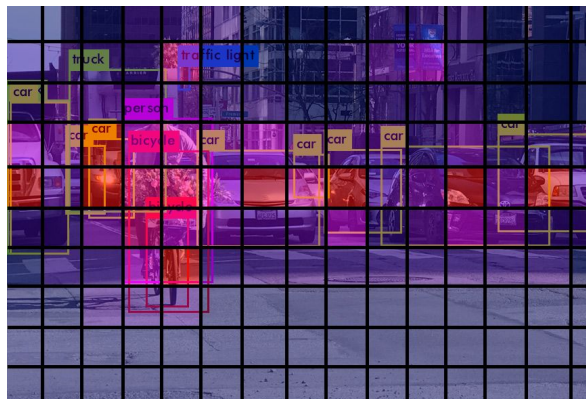
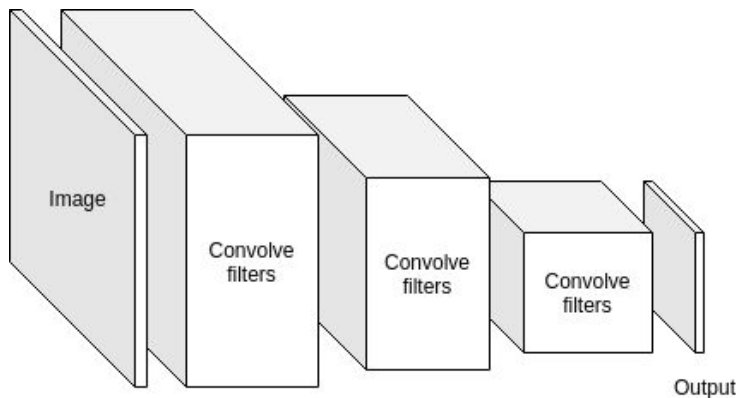
Detection

- We detect lines of text/writing using an object detection method.
- Specifically we model our detector after YOLOv2/v3.
 - Single fully convolutional network



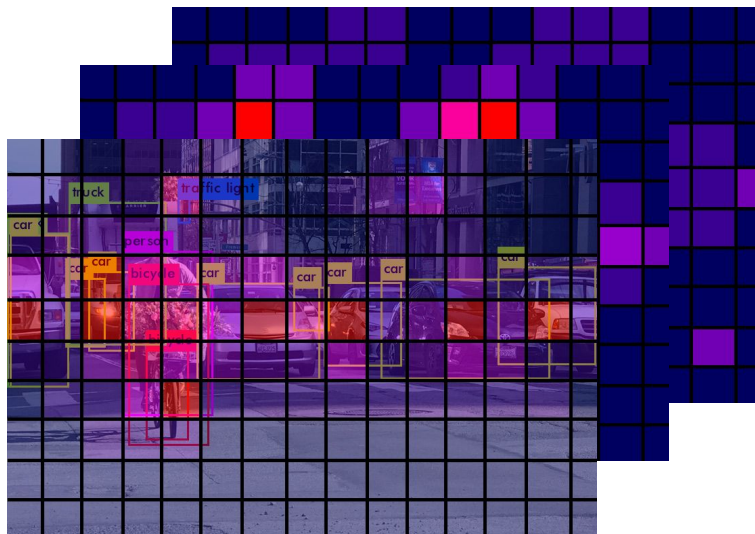
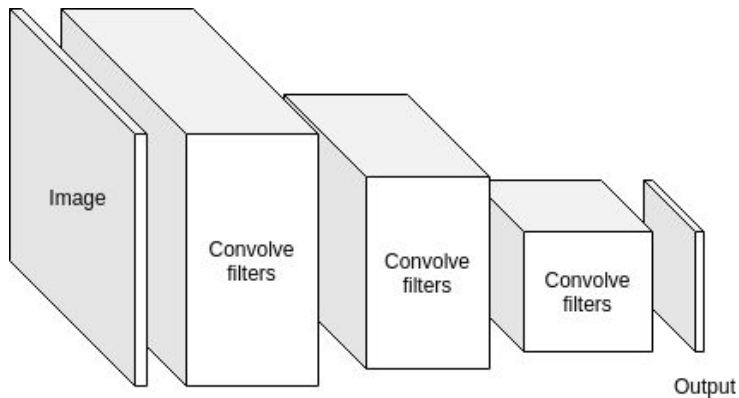
Detection

- We detect lines of text/writing using an object detection method.
- Specifically we model our detector after YOLOv2/v3.
 - Single fully convolutional network



Detection

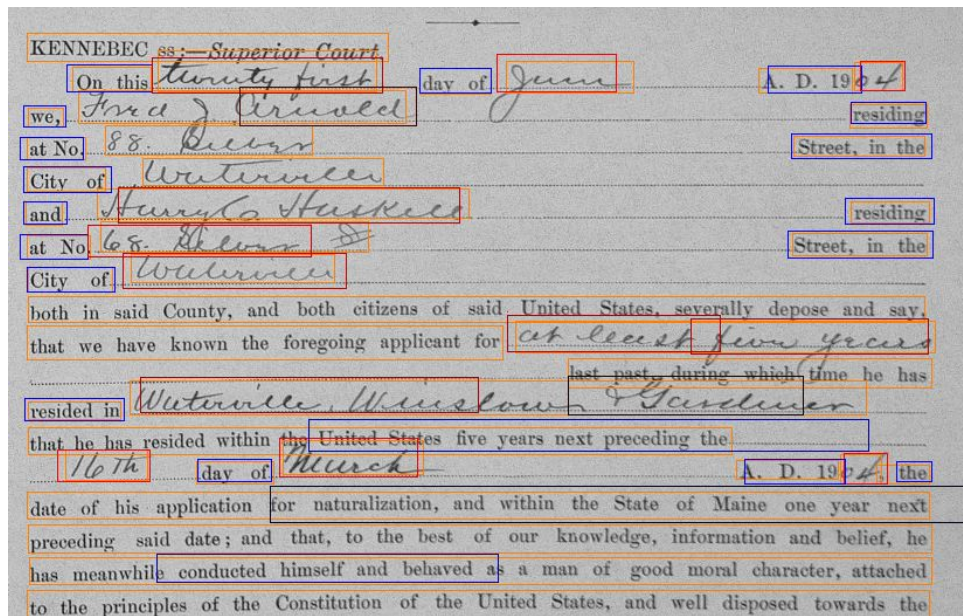
- We detect lines of text/writing using an object detection method.
- Specifically we model our detector after YOLOv2/v3.
 - Single fully convolutional network



Detection

- We detect lines of text/writing using an object detection method.
- Specifically we model our detector after YOLOv2/v3.
 - Single fully convolutional network

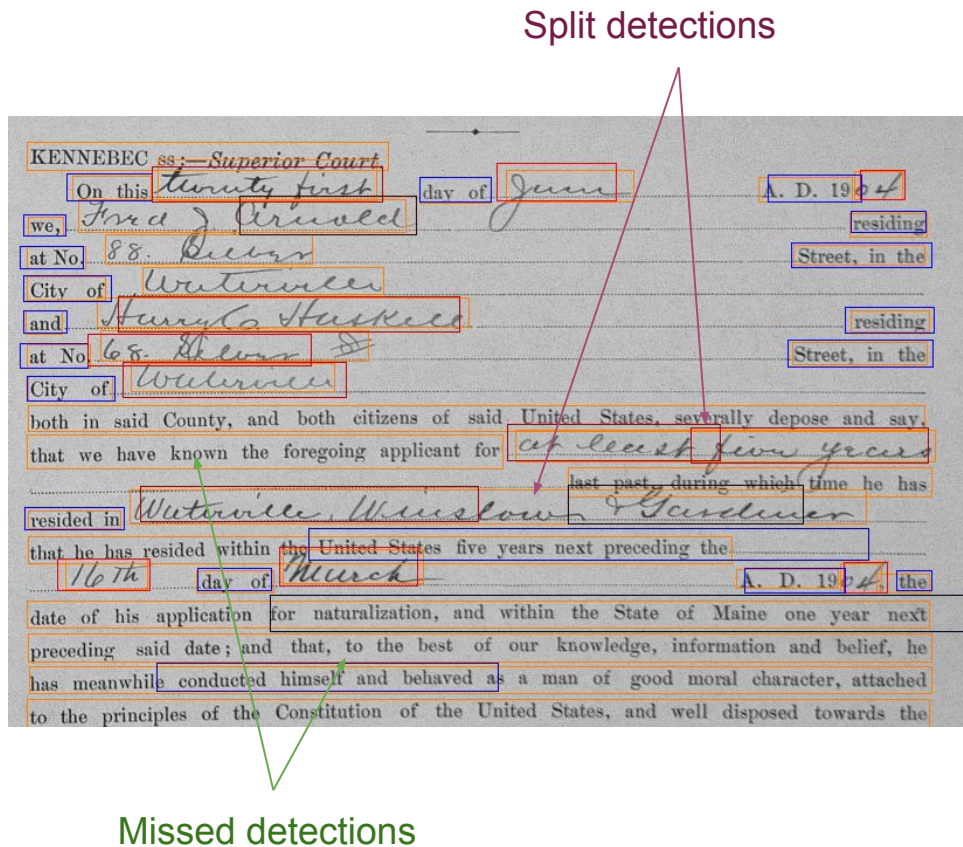
However, a normal convolutional net architecture does not handle very long text/writing lines.



Detection

- We detect lines of text/writing using an object detection method.
- Specifically we model our detector after YOLOv2/v3.
 - Single fully convolutional network

However, a normal convolutional net architecture does not handle very long text/writing lines.

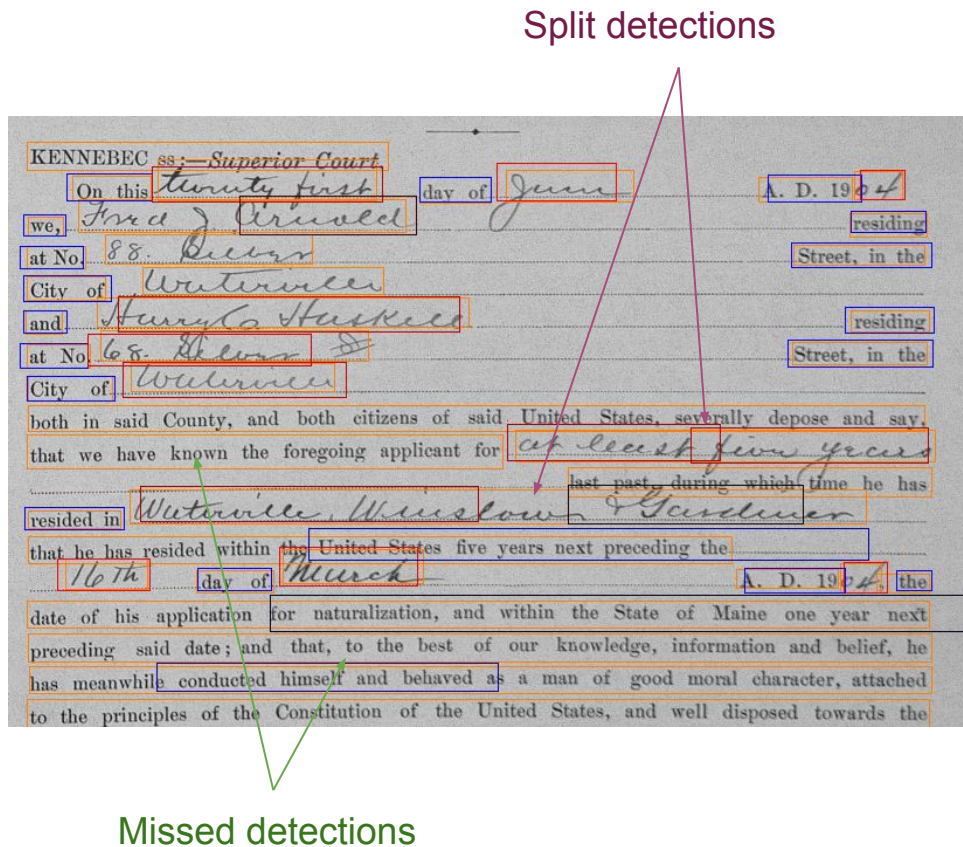


Detection

- We detect lines of text/writing using an object detection method.
- Specifically we model our detector after YOLOv2/v3.
 - Single fully convolutional network

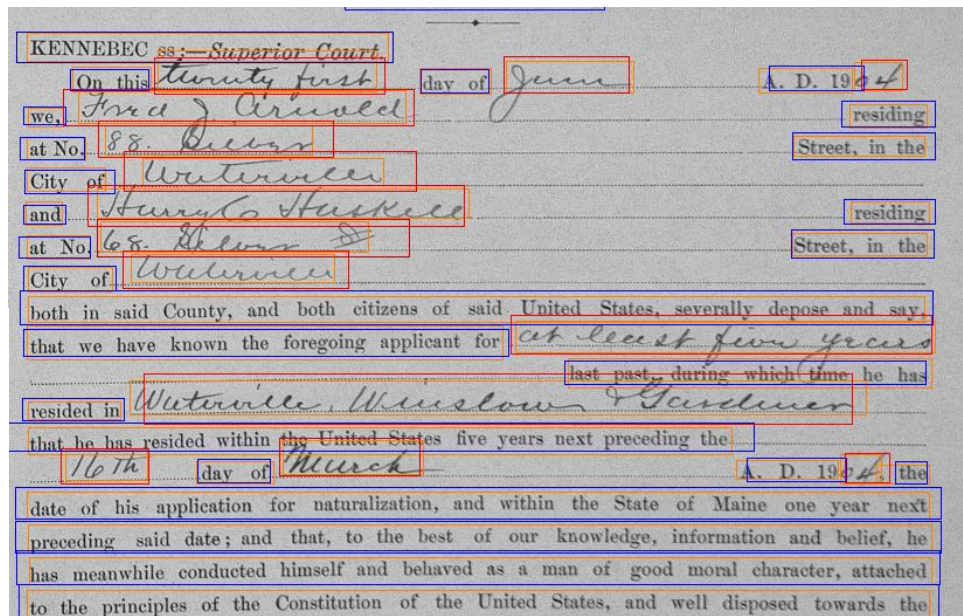
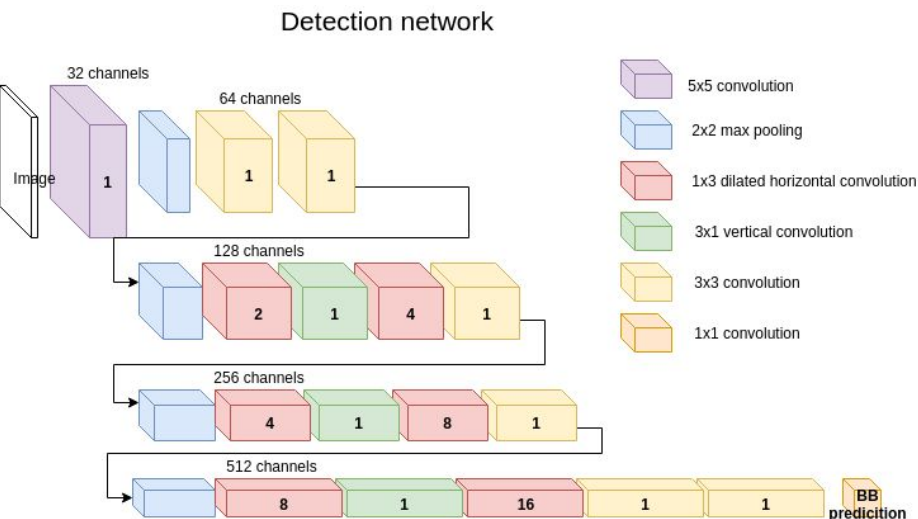
However, a normal convolutional net architecture does not handle very long text/writing lines.

YOLO handles this with multi-resolution detection.



Detection

We increase the horizontal receptive field of our network by introducing dilated 1x3 convolutions.



Detection Results

Mean average precision (measures both precision and recall)

Normal convolutional network: 0.397 mAP

With dilations (our architecture): 0.453 mAP

I, Jos H. Miceuth, Seaman that the within List
contains the Names of the Crew of the Brig Salem
together with the
Place of their Birth and Residence, as far as I can ascertain the same.
Wm this 18th day of November 1834
Before me, W. M. Ward Collector

(Signed)
Jos H. Miceuth

Brig Salem
Sept. 16th
1835

1834

I do certify, that the within is a true Copy of the List of the Crew of the Brig Salem
at Salem whereof A. H. Miceuth
is Master, taken from the Original, on File in this office.

Given under my hand and Seal of Office, at the Customs-House,
this eighteenth day of November
in the Year of our Lord one thousand eight hundred and thirty four

W. M. Ward
Collector

I do hereby further certify, That the within named
Nel W. Andrews
Sam^r West 3^d
Charles Booth
Sam^r Blythe
James Kenney
George Taylor

Examined by W. M. Ward Collector

I do hereby certify, that the said
Persons
above named
3
are citizens of the United States of America.
Given under my hand and Seal of Office, this 18th day of Novr 1834

W. M. Ward
Collector

NOTE PARTICULARLY.—For use of Aliens who arrived BEFORE June 29, 1906.

FACTS FOR PETITION FOR NATURALIZATION

Clerk of court and applicant should read these instructions carefully.
Clerk of court should refer to recent petition for aliens who arrived in the United States after June 29, 1906, until certificate of arrival is furnished the clerk by the Bureau of Immigration. Your request for the name or new name "must show" the date of the registration.
A copy of this form (216) should be furnished by the clerk of court to each applicant for a petition for naturalization who arrived in this country on or before June 29, 1906, so that he can fill in the spaces in the certificate.
The clerk of court must collect the fee of \$5 before he commences to fill out the petition.
For an explanation of each block or for the questions thereon to the clerk or deputy clerk.
After being filled out, the form is to be returned to the clerk, to be used by him in properly filling out the petition.
Witnesses must be citizens of the United States. If any witness is a substituted citizen, he must take the certificate of citizenship or recalled record paper, to the court when the petition is filed, and also when it is read by the court.

My name is Robert Stobie also known as Robert Stobie
1. My place of residence is 121 1/2 St. Lewiston
2. My present occupation is tailor
3. I was born on the 20 day of October 1875 at Pruscha, Province, Russia
4. I emigrated to the United States from Pruscha, Province, Russia on or about the day of August, 1895, and arrived in the United States at the port of Salem, Mass. on the 10 day of September, 1905, on the vessel Swansea
5. I declared my intention No. 1745, to become a citizen of the United States on the 15 day of November, 1897, at _____ in the _____ Court of _____

6. I am married. My wife's name is _____. She was born in _____ on the _____ day of _____, 18_____, and now resides at _____. I have children, and the name, date, and place of birth, and place of residence of each of said children is as follows:

| born | day of | at | resides at |
|------|--------|----|------------|
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |
| | | | |

7. I now owe allegiance to _____.
8. I am able to speak the English language.
9. I have resided continuously in the United States since the 10 day of Sept, 1895, and in the State of ORADO since the 15 day of Sept, 1905.
10. I have heretofore made petition for United States citizenship.

If petitioner has heretofore made application for citizenship, the facts required should be fully stated in the following blanks:
I previously petitioned for citizenship to the Court, at _____, on the _____ day of _____, 1_____, which was denied for the following reason: _____

The cause of such denial has since been cured or removed.

How names, addresses, and residence addresses of witnesses, citizens of the United States, who have known you for at least five years, but not all members of one family, in which petition is made, who will make affidavit that you are a person of good moral character, who you are qualified to receive the citizenship of the United States, and who will appear with you before the court, when petition is filed, and to the person of the filing of the petition.

James Bragantini Shipping Clerk residing at
James E. Burke Lewiston 179 E 7 1/2 St. Lew.
Signatures of witnesses who will be introduced by me if there appears with me at the time of filing my petition for naturalization are made to appear at the time of hearing.

Each of the above-named witnesses has known me at least _____ years.
I herewith present my Declaration of Intention to become a citizen of the United States.

11A

Wherefore, your applicant prays, that he may be admitted to become a citizen of the said United States of America, according to the forms of the statutes in such case made and provided.

June 21 1904 Sworn to by the above-named applicant before the Court.
Attest: Le. M. Jones Clerk.

State of Maine.

Kennebec Superior Court.
On this twenty first day of June, A. D. 1904.
I, Robert Stobie, do solemnly swear and declare on oath that I will support the Constitution of the United States, and that I will and do absolutely and entirely renounce and adjure all allegiance and fidelity to every foreign Prince, Potentate, State or Sovereignty,—particularly to Edward VII., King of the United Kingdom of Great Britain and Ireland, Swegen Sweden.

whoso subject I have heretofore been,—so help me God.
June 21 1904 Sworn to by the said Robert Stobie before the Court.
Attest: Le. M. Jones Clerk.

State of Maine.

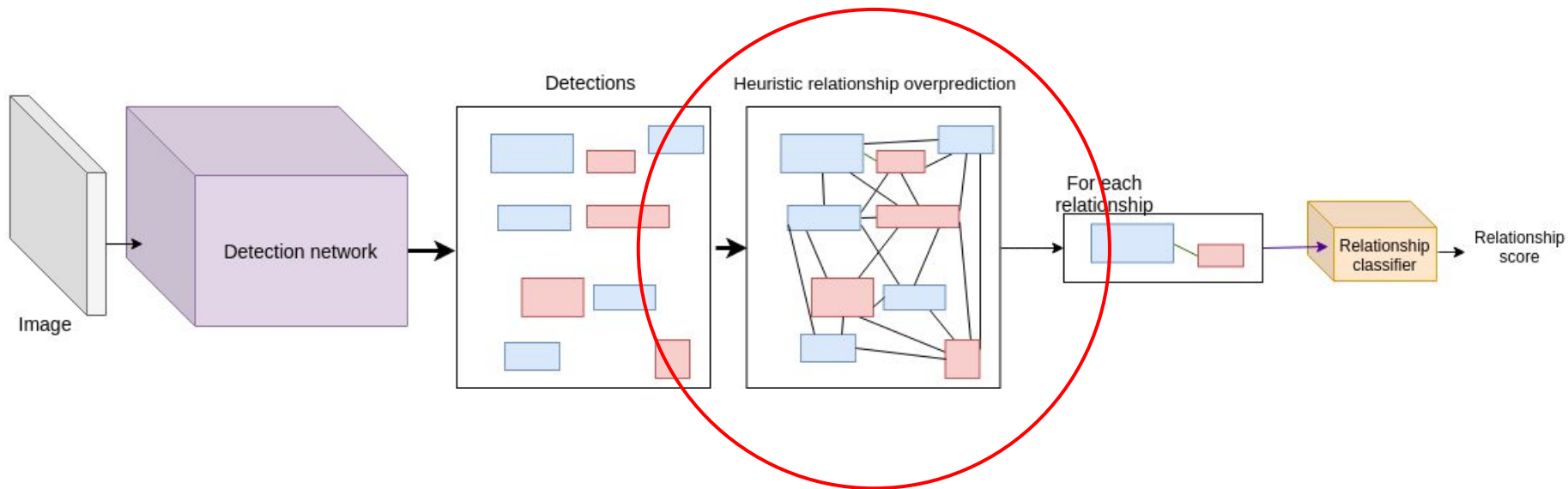
Kennebec Superior Court.
On this twenty first day of June, A. D. 1904.
I, James E. Burke, residing _____ Street, in the _____ City of WATERVILLE and Henry H. Haskins residing _____ Street, in the _____ City of WATERVILLE

both in said County, and both citizens of said United States, severally depose and say that we have known the foregoing applicant for at least five years last past, during which time he has resided in WATERVILLE, MEADOWS & GARDNER that he has resided within the United States five years next preceding the 10th day of March, A. D. 1904, the date of his application for naturalization, and within the State of Maine one year next preceding said date; and that, to the best of our knowledge, information and belief, he has meanwhile conducted himself and behaved as a man of good moral character, attached to the principles of the Constitution of the United States, and well disposed towards the good order and happiness of the same, and obeyed the laws thereof.

June 21 1904 Sworn to by the above-named Witnesses, before the Court.
Attest: Le. M. Jones Clerk.

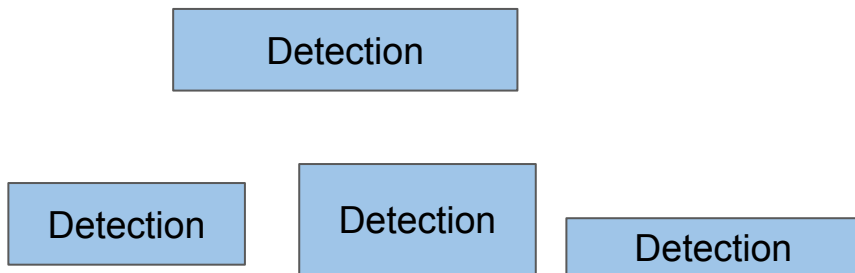
SUPERIOR COURT KENNEBEC CO WATERVILLE ME VOL. 10 PAGE

Method



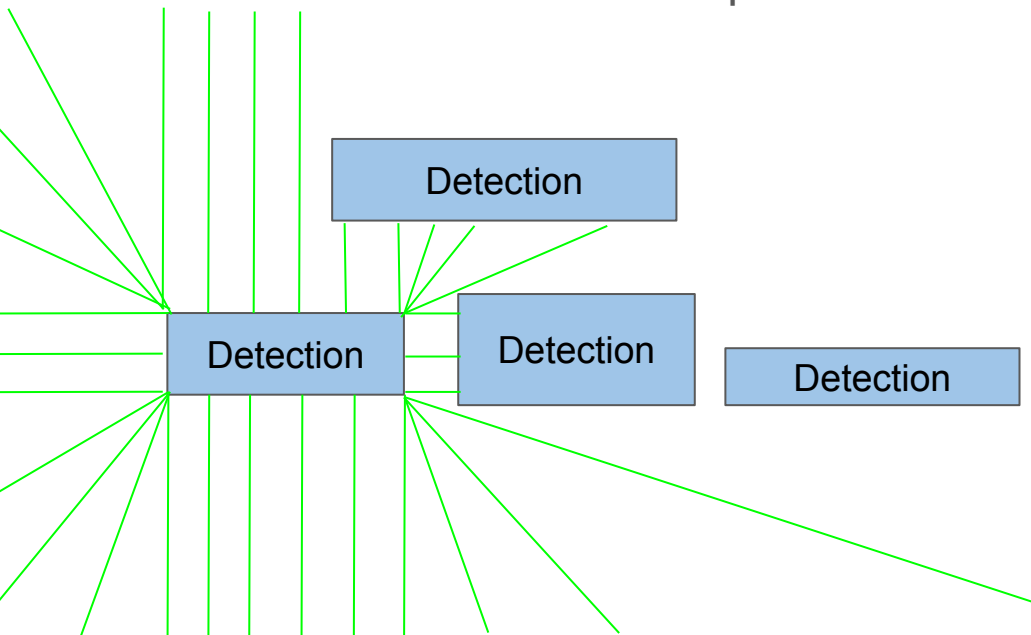
Heuristic over pairing

- Rather than considering every possible pairing, we only consider a pair between detections that are within line-of-sight.
- +0.99 recall on relationships



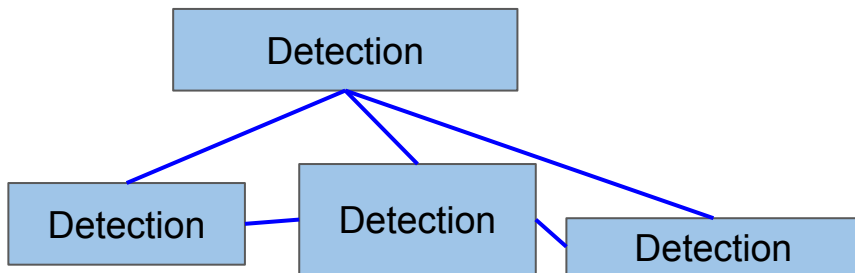
Heuristic over pairing

- Rather than considering every possible pairing, we only consider a pair between detections that are within line-of-sight.
- +0.99 recall on relationships

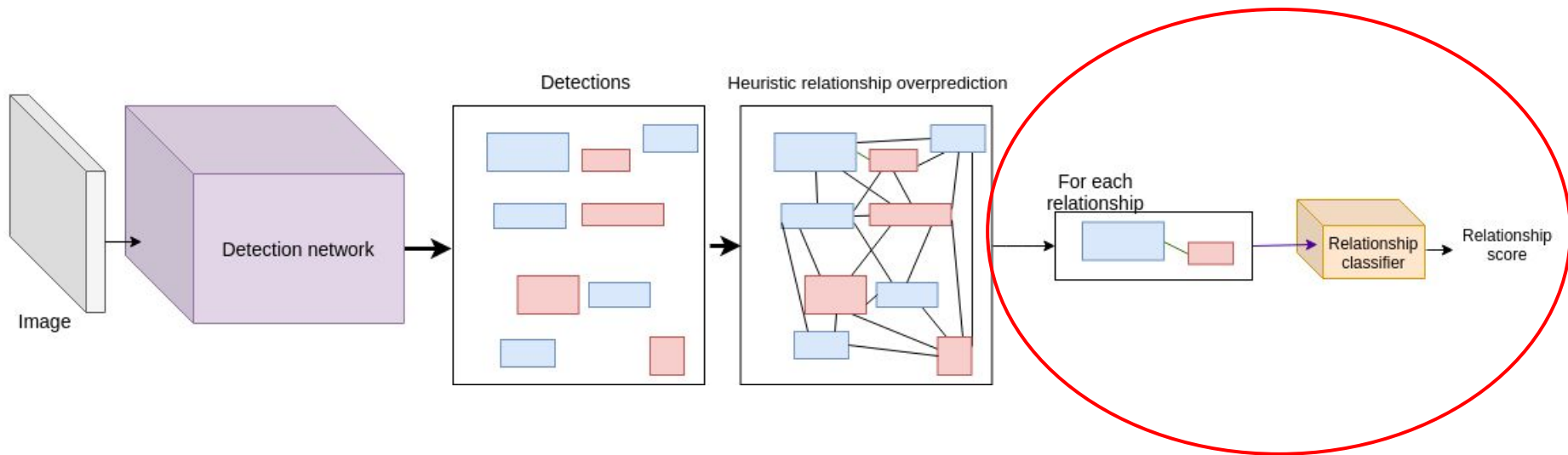


Heuristic over pairing

- Rather than considering every possible pairing, we only consider a pair between detections that are within line-of-sight.
- +0.99 recall on relationships



Method



How to determine relationships?

- Nearest neighbor of opposite class
 - Simple, but definitely not optimal

How to determine relationships?

- Nearest neighbor of opposite class
- Heuristic scoring functions (borrowed from Hirayama et al. “Development of template-free form recognition system”)
 - Average of three functions
 - It's too the left
 - Same height
 - Not too far apart
 - Better, but still not optimal

How to determine relationships?

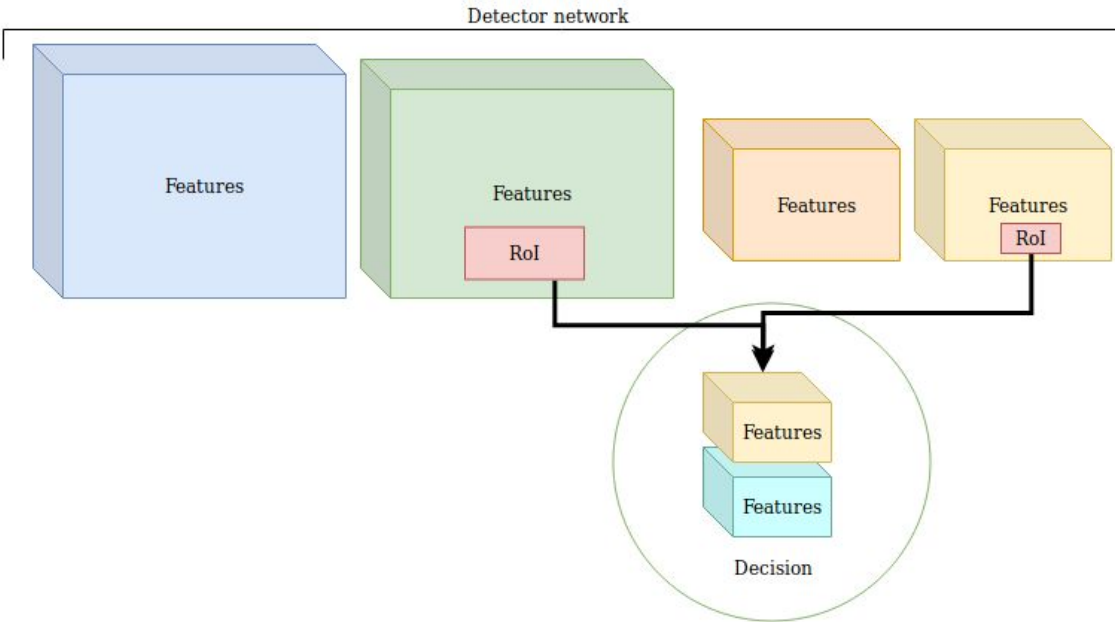
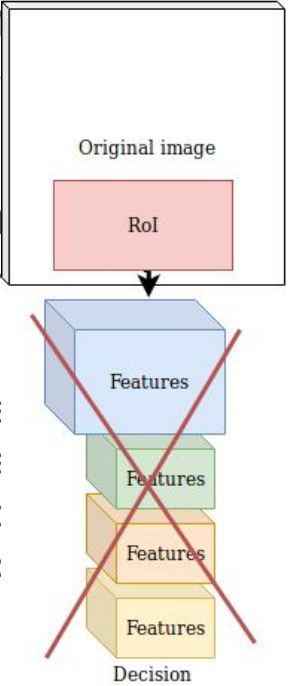
- Nearest neighbor of opposite class
- Heuristic scoring functions(borrowed from Hirayama et al. “Development of template-free form recognition system”)
- Why not learn heuristic rules? Give shapes and distances as features to machine learning classifier (neural network)
 - Heights, widths, x-offset, y-offset, distances of each corner
 - Two hidden fully connected layers, ReLU, batch norm
 - Trained from detector output

How to determine relationships?

- Nearest neighbor of opposite class
- Heuristic scoring functions(borrowed from Hirayama et al. “Development of template-free form recognition system”)
- Why not learn heuristic rules? Give shapes and distances as features to machine learning classifier (neural network)
- Let us see!
 - Why not use a convolutional network look at the potential relationships?
 - RoIAlign features from detector network (reusing features)
 - Small convolutional network trained with binary cross entropy
 - Detector is pre-trained, then fine-tuned as classifying network trains.

How to determine relationships?

- Nearest neighbor
- Heuristic score
- template-free
- Why not learn machine learning
- Let us see!
 - Why not use
 - RoIAlign fea
 - Small conv
 - Detector is p



Optimization

Pairing decisions should depend on each other.

We frame this as an optimization over all relationships, given a prior of how many neighbors/relationships each detection should have

Venenatis tellus in metus vulputate

Cras semper auctor neque vitae tempus quam pellentesque nec nam. Adipiscing elit duis tristique sollicitudin nibh sit amet. Faucibus a pellentesque sit amet porttitor eget dolor morbi. Ipsum suspendisse ultrices gravida dictum fusce.

Ut enim:

>Lorem ipsum dolor

ad minim:

Exception

Veniam:

Quis nostrud:

Cras sed felis

Semper:

Exception

Feugiat nibh:

Velit aliquet

Pulvinar proin:

1998

Gravida hendrerit:

Nulla 25th

Voluptat:

Vivamus at augue eget arcu dictum varius duis at. Sit amet dictum sit amet justo donec enim diam. Id donec ultrices tincidunt arcu non sodales neque sodales. Mi proin sed libero enim. Sit amet cursus sit amet dictum sit amet justo donec. Sodales neque sodales ut etiam. Morbi leo urna molestie at elementum eu facilisis. Tristique sollicitudin nibh sit amet commodo nulla. Sed egestas egestas fringilla phasellus faucibus scelerisque eleifend. Risus pretium quam vulputate dignissim. In ante metus dictum at tempor commodo ullamcorper a. Sed risus ultricies tristique nulla aliquet enim tortor at. Duis ut diam quam nulla. Auctor eu augue ut lectus. Ullamcorper morbi tincidunt ornare massa eget. Platea dictumst quisque sagittis purus sit amet volutpat consequat.

Lacinia Quis

Viverra

Hendrerit Gravida

Sed egestas

Vestibulum

Quis justo

Sed egestas

Feugiat

Sed egestas

Optimization

Intuition: We want to pick relationships (x_r) with high probability ($p_r > 0.5$), but we don't want more relationships (sum x_r) than we think a text instance has (n_i).

c is a constant indicating how much we trust n_i

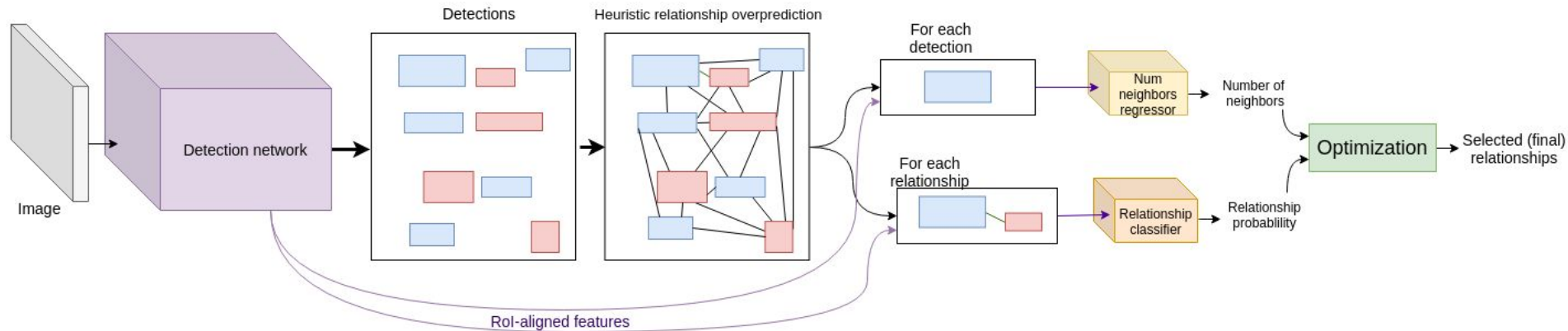
$$\begin{aligned} & \underset{X}{\text{maximize}} && \sum_{r=1}^R (2p_r - 1)x_r - c \sum_i^B (n_i - \sum_{x_r \in R_i} x_r)^2 \\ & \text{subject to} && x_r \in \{0, 1\}, \end{aligned}$$

Optimization

We need a prediction of the number of neighbors (generally 0 or 1)

- We can have the detector network predict the number of neighbors (along with class, etc.)
- We can have an additional convolutional network (parallel to relationship classifier) predict the number of neighbors

Method



Results for Label-Value pair finding

| | mAP | mAP with optimization |
|---------------------|-------|-----------------------|
| Nearest neighbor | 0.197 | 0.243 |
| Heuristic rules | 0.244 | 0.243 |
| Blind network | 0.325 | 0.349 |
| Full/visual network | 0.566 | 0.566 |

True positive occurs when both detections have >0.5 IoU with ground truth bounding boxes and there is a true relationship between them.

MEDICAL CERTIFICATE.

Port of Baton Rouge 1919
 Name, Olsen Knud
 Age, 33 Sex, M
 Nat., Norway Race, 55
 Date arrival, 11-11-19
 S. S. Wassinda
 Class, 1st Manifest No.

This is to certify that the above-described person has this day been examined and is found to be afflicted with:

Sacculliciae
Arthritis, may
require treatment
may affect ability
& earn a living.

Surgeon.

Surgeon.

Surgeon.

Admitted U.S.
 PRIMARY INSPECTION MEMORANDUM
 NEW YORK

Department of Commerce and Labor

U. S. IMMIGRATION SERVICE

Date of arrival 11-11-19
 Name Olsen Knud
 Age 33 Sex M Married single
 Occupation Teacher
 Able to read Yes Write Yes
 Nationality Norwegian
 Race White
 Last permanent residence Norway
 Name and complete address of nearest relative or friend in country of origin name Wassinda
 Address 14th Street New York
 Destination United States
 Ticket Yes In his possession Yes
 Amount of money 19.66 Dollars
 Ever in U.S. Yes 1916
 Date of last departure from U.S. 1916
 Going to join whom? Friend
 Relative Name Wassinda
 Address 14th Street New York
 Whether under contract No
 Condition of health Good
 Landing: Yes
 Port New York
 Date 11-11-19
 Steamship Wassinda

New York Branch Freedmans Union Commission.

TEACHER'S SCHOOL REPORT.

For the Month of May 1869

Name and Location of school, "Barnwick Co" - Kammerville, Va.
 When opened, September 25th 1868 Inclusive June 30th 1869
 Number of Teachers, Four White, Two Colored, 0 Northern, Two
 By whom the Teachers are employed, By Board of Freedmans Union Commission
 Number of Days kept, Twenty Sessions per day, Two Length of Session, 6 1/2 hours
 Number of Pupils enrolled, 97 Boys, 42 Girls, 55 White, 1
 Number of Pupils enrolled as free school, One hundred & two (102)
 Number left School this month, Eight (8)
 Number of new Pupils this month, Three (3)
 Number always present, 34 Always Punctual, 26 Average Attendance, 66
 Number over 16 years of age, 10 Under six, 2
 Number of Pupils receiving tuition, 28 Amount paid to you this month, \$7.36
 Number in Reader and Primer, 16 In 1st of Series, 21 2d, 16 3d, 20 4th, 20 5th, 1
 Number in Mental-Arithmetic, (81) Eighty one
 Number in Written Arithmetic, 60 In 2d of a Series, 6 In 3d, 0
 Number in Geography, 61 In 2d of a Series, 14 In 3d, 8
 Number in Grammar, (44) Forty Four
 Number who write on slates, 61 In Books, 20
 Number at Needlework, 0
 Is your School master? Yes How many grades? Two
 Have you a Sunday-School? Yes Number of Pupils, 112 Of Teachers, 6
 Have you a Class in Sunday-School? Yes Number of Pupils, 35
 Have you a Bible-chorus? Yes Number of Pupils, 7 Of Teachers, 1
 Have you a Class in English-grammar? Yes Number of Pupils, 0

Each blank should be filled with a number or cipher, and the Report forwarded as soon after the first of the Month as possible. Remarks, in letter form, to be made by the teacher on the third page of this sheet.

Conclusion

- Trainable method for detecting pre-printed text and input writing
- Trainable method for determining relationships
 - Outperforms heuristic baselines
- Method is language agnostic
 - Reading text may help in pairing
- This is one piece of a bigger problem

Thank you!

Questions?