

Image Analysis ECSS projects update



Decomposing Bodies (PI A. Langmead (Univ of Pittsburgh):

~20K early 20th century
Bertillon prison id cards

analyzing, digitizing and re-presenting
the data

examine information management of
decomposing bodies into a series of
numerical and visual components.

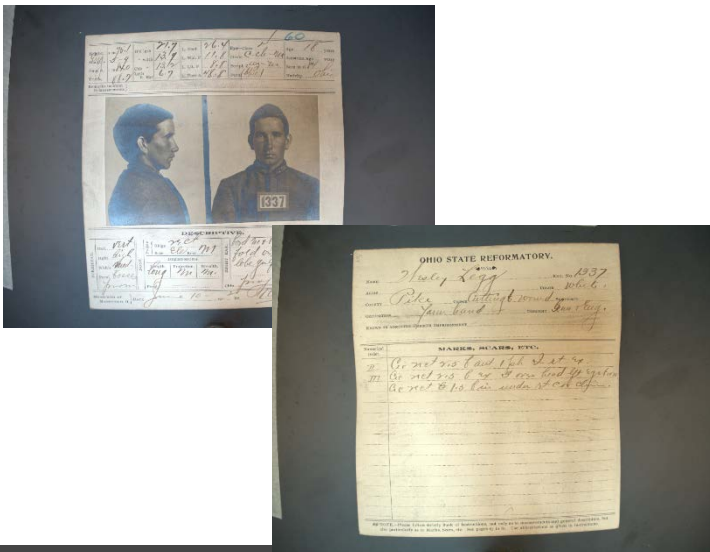


Image Analysis of Rural Photography (PI L. Wuerffel, J. Will, Valpraiso):

~200K depression & WWII era photos

visual documentation and visual rhetoric

analysis of image content

analysis of Image characteristics

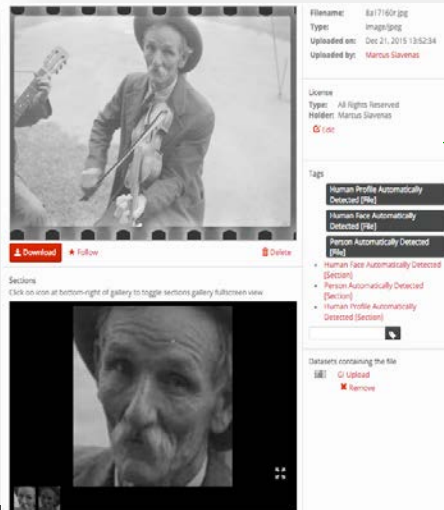
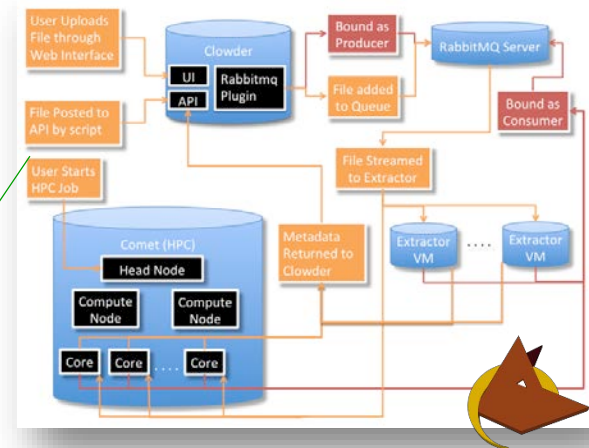


Infrastructure Support for Data Mining – using NCSA BrownDog and Clowder interface



Make Files Searchable

Run Feature Extractors on COMET.



Search Files in Clowder with Feature Tags

Marcus Slavenas, Sandeep Satheesan NCSA
Paul Rodriguez, SDSC
Alan Craig, XSEDE
Elizabeth Wuerffel, Valparaíso University
Jeffrey Will, Valparaíso University



Extracting features for datamining

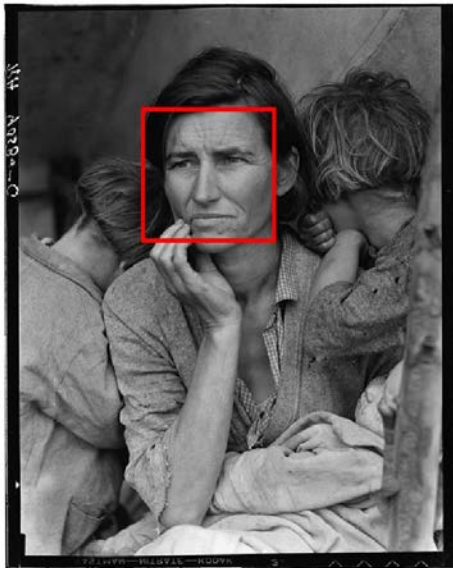
Visual Content: *Faces, Eyes, Profiles (openCV models)*
Visual Characteristics: *Hole punched photos, gray scale*
Metadata: *Geo-location, Photographer, Date, Subject*
Title Content: *Word part of speech and semantics*



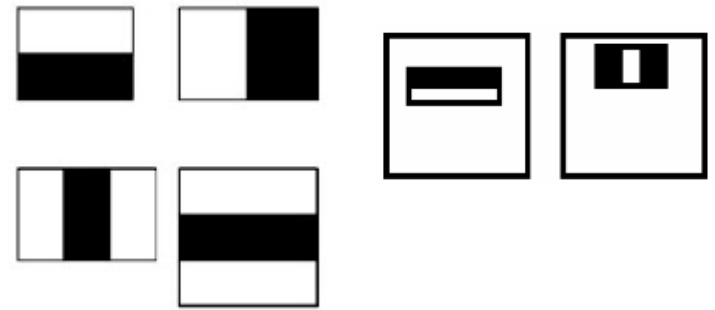
Creator: Lange

Title:

"Destitute pea pickers in California.
Mother of seven children."



x



*OpenCV Face
Detector:*

*Apply each filter to
find darker/lighter
patterns*

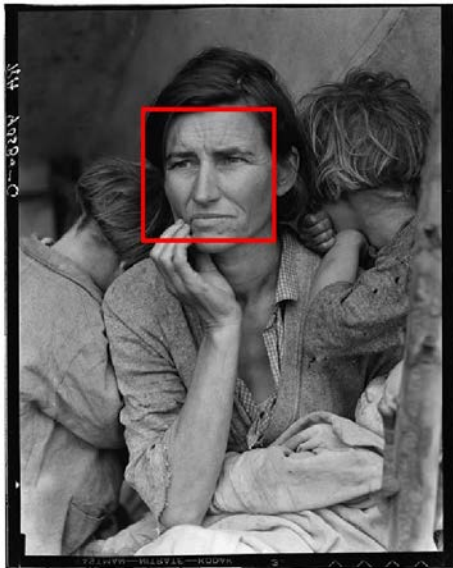
*Each filter highlight
facial characteristics*

*Train classifier on
many samples*

Creator: Lange

Title:

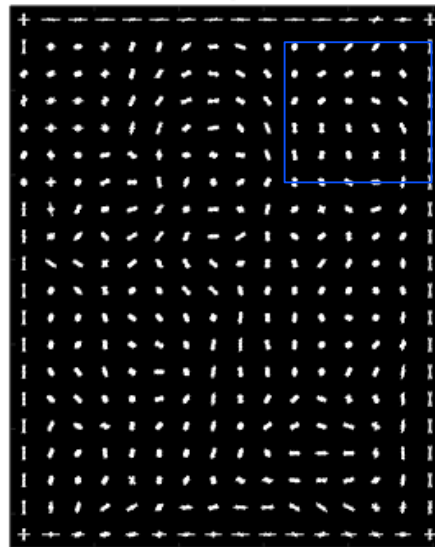
”Destitute pea pickers in California.
Mother of seven children.”



Histogram of Gradients

CellSize = [32 32]

Feature length = 10260



*OpenCV: gradient
information*

*For each pixel in a
subregion, get
average contrast
gradient*

*Useful for people
identification*

Example title:

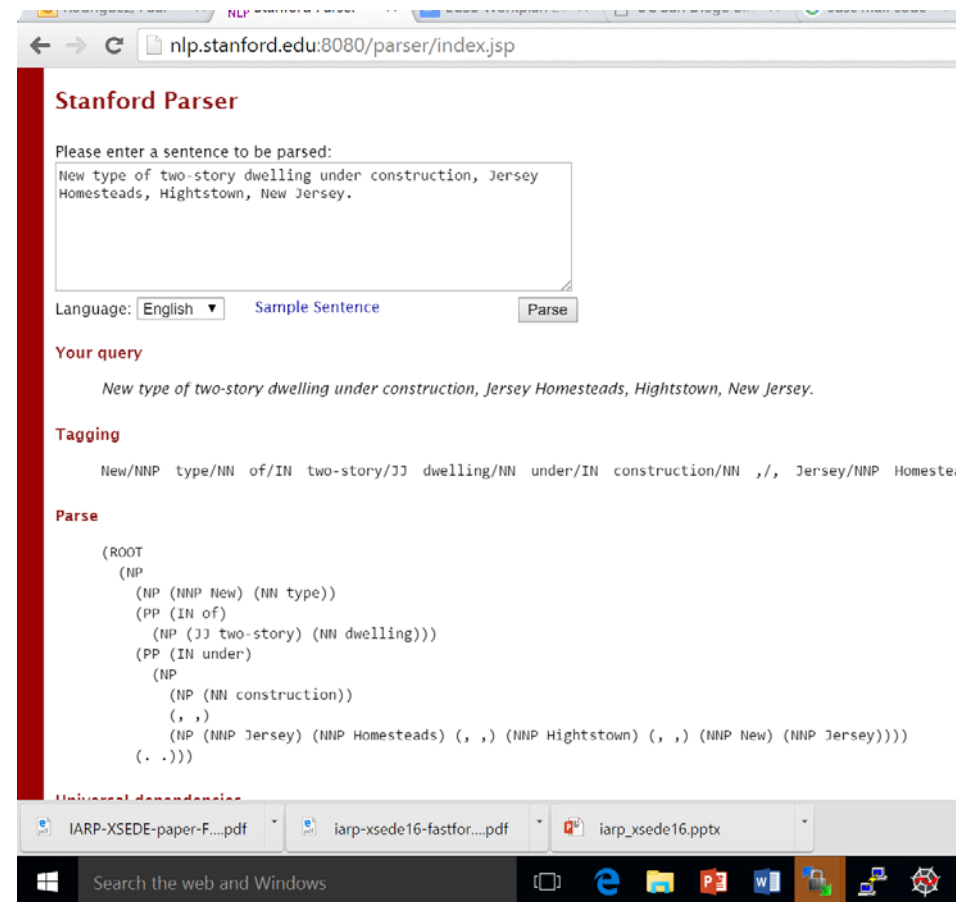


Image 20950: 'New type of two-story dwelling under construction, Jersey Homesteads, Hightstown, New Jersey.'

- **Develop python batch pipeline to extract lexical-semantic features: named entities, part of speech, semantic category. About 5hours for 5000 titles on 1 COMET node.**

Using Stanford NLP tools, Python NLTK (nat. lang. tool kit)

- Extract & Replace name entities (otherwise Mr. Smith might be a city)
- Extract & Replace city-states (to make parsing easier)
- Parse and get part of speech
- Search ontology for each Noun
- Find Nouns that possibly belongs to category of interest



Using Wordnet semantic ontology

- Synomyn sets (ie search first 2 Noun sets)
- Check lemmas (ie only use words with 3 common letters; eg pop not~ dad
dad ~ daddy)
- Look at upto 9 semantic levels for:
 - person
 - structure
 - artifact
 - animal
 - object
 - physicalentity
 - road
 - place

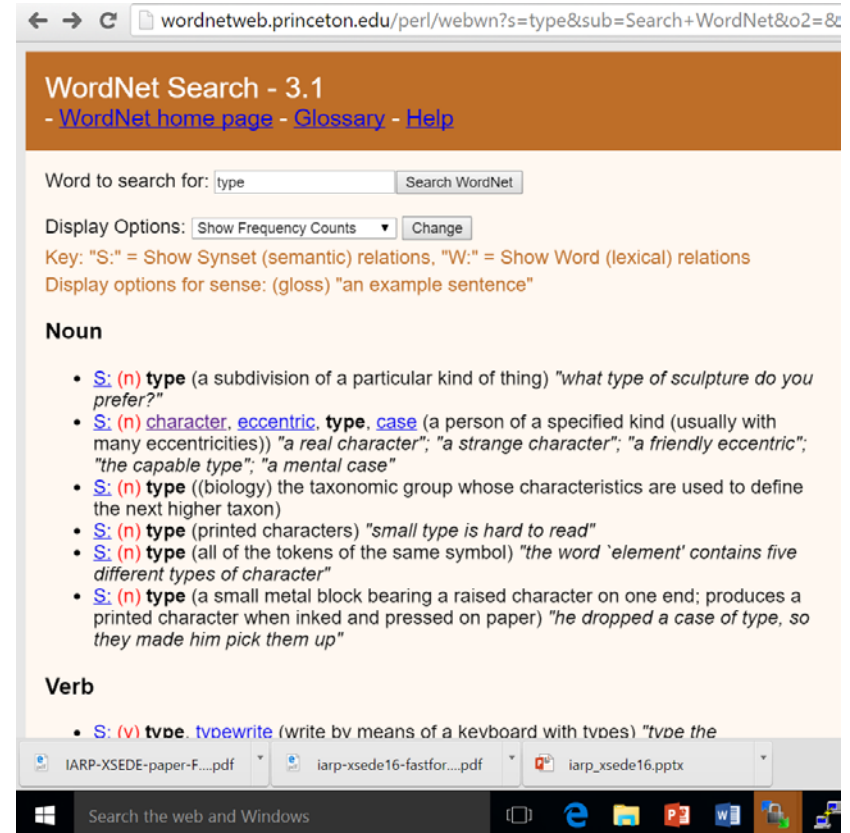
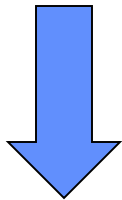
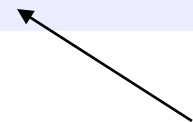


Image 20950: 'New type of two-story dwelling under construction, Jersey Homesteads, Hightstown, New Jersey.'



Word/Features Image 20950	Part of Speech	Semantic Category	Relative Frequency (of this sense)	Depth in Parse Tree
'type'	Noun	Person >> Agent	Low	1
'dwelling'	Noun	Housing>>Structure	High	1
Jersey Homesteads	Named Entity	n/a	n/a	1
'construction'	Noun	n/a	n/a	0



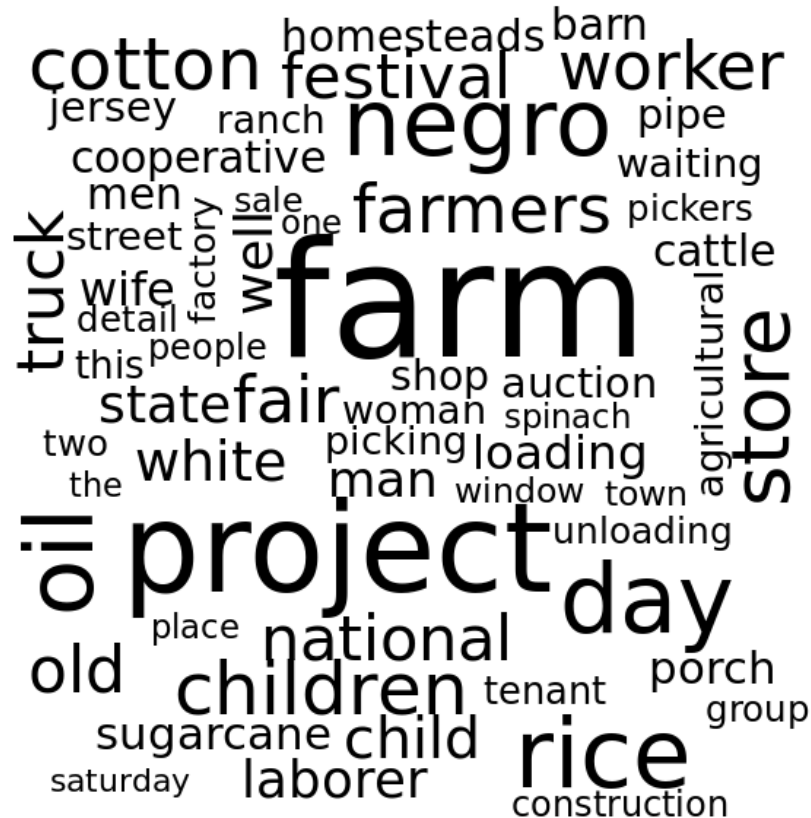
Abstract words are not processed for semantics

Cleaning up titles allows word clouds

A word cloud for 5000 image
sample.

Locations, named entities,
abbreviations, small words
removed.

Text uploaded to Text
Analytics Gateway



Example Query

“Select all images where there is a word in the semantic category ‘animal’ .”

About 7% (275 out of 5000) have some mention of an animal, or animal related topic, with 73 different subject categories, such as ‘Farms,’ ‘Auctions,’ ‘Small towns,’ ‘Spinach workers,’ and so on.

False positives: ‘YOUNG’ possibly refers to <animal>

Example Query

- *Select all pictures by Lange with possible 'person' and num_faces > 0*
- How to view results that are image items?

1935_2_46439_faces_Children_of_Oklahoma_droug
ht_r

1935_2_46442_faces_Jack_Neill,mig
ratory_fruit_tra

1935_5_46442_faces_Jack_Neill,mig
ratory_fruit_tra

1936_1_49064_faces_Destitute_pea
_pickers_in_Calif

1936_2_49109_faces_Accepted_ap
plicant_for_resettl

1936_2_49116_faces_Cherry_picker
s_near_Millville,

1936_2_49153_faces_That_distinct_
American_frontie

1936_3_49188_faces_Children_of_t
urpentine_worker

1936_3_49220_faces_Children_of_e
victed_sharecrop
p

1936_3_49247_faces_Children_at_H
ill_House,Mississ

1936_3_49290_faces_Child_living_i
n_Oklahoma_City

1936_3_49326_faces_Child_of_an_i
mpoverished_fa
mil

1936_3_54227_faces_Thomaston,(
vicinity),Georgia__

1936_4_49429_faces_Children_of_d
estitute_family__

1936_4_49497_faces_Children_and
_home_of_migrat
ory

1936_11_49497_faces_Children_and
d_home_of_migr
atory

1937_1_55849_faces_Japanese_mo
ther_and_daught
er,a

1937_1_55939_faces_This_is_a_ha
rd_way_to_serve_t
h

1937_1_55987_faces_Children_of
migrant_agricultu
r

1937_1_56007_faces_The_Great_Re
aping_Day__Hym
n_si

1937_2_56047_faces_Children_of_

1937_2_56093_faces_Children_fro

1937_2_56183_faces_Children_of

1937_2_56185_faces_Children_of

1937_2_56290_faces_Child_of_for

1937_2_56313_faces_This_man_wa

1937_2_56346_faces_Child_of_Texa

1937_2_56367_faces_Child_of_shar

1937_2_56400_faces_Delta_cooper

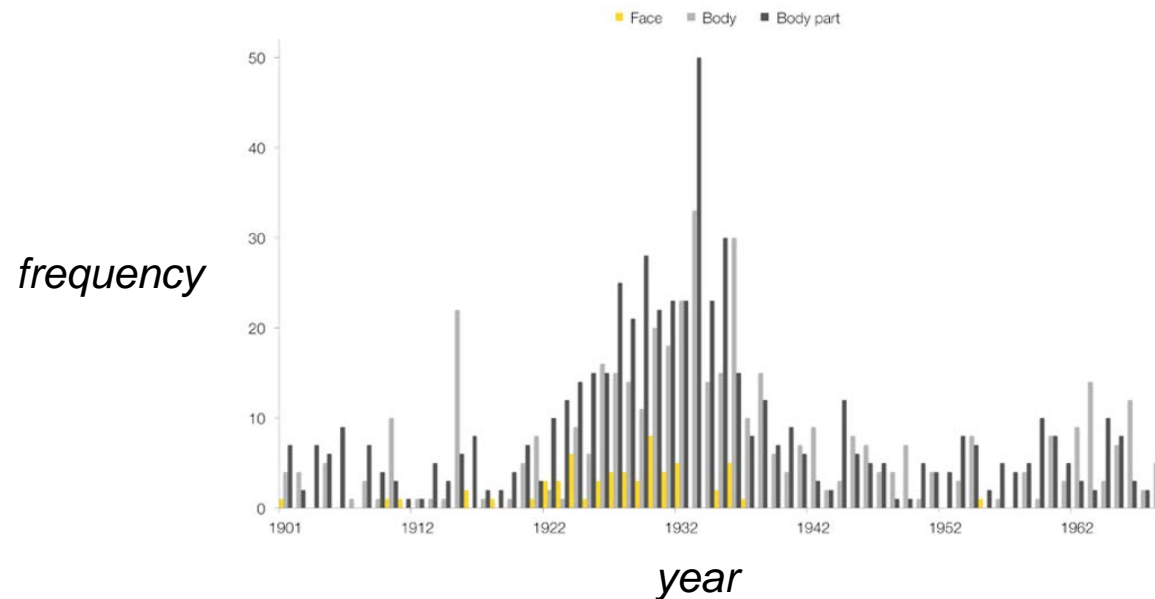
1937_2_56438_faces_Children_of_t

Use SQLite3 on COMET to process tables

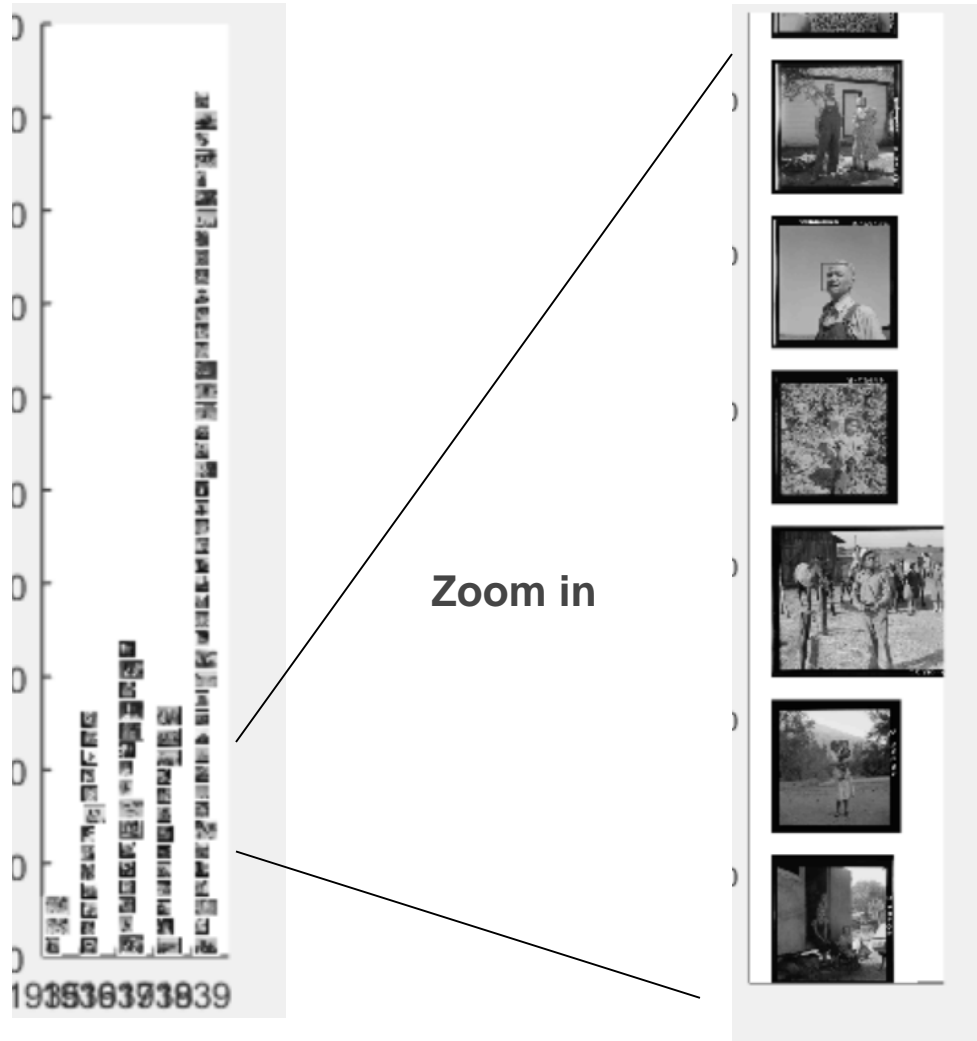
Use Matlab to gather file names and draw bounding boxes on faces

Scatter plots and histogram with thumbnails from Lev Manovich

(using art collection of photographs, **manually tagged** for faces and bodies)



Perhaps view by a scatter plot with images (ie Manovich)



Decomposing Bodies

Sandeep Satheesan NCSA

Paul Rodriguez, SDSC

Alan Craig, XSEDE

**Alison Langmead, Univ of
Pittsburgh**

- **From the ECSS request:** *Decomposing Bodies* seeks to de-familiarize this process of breaking down and defining what we see into quantized digests, by collecting, analyzing, digitizing and re-presenting the data created by the process of Bertillonage.



- **From Pis digital media project (A. Langmead):** *Data (after)* *Lives* investigates the variegated relationship between human notions of the self and these procedures that produce alternative, externalized, malleable representations of the human experience.

Front and back of prison cards - some smudges, warping, noise. What characters can we recognize?

76.1 175 137 137 67		26.4 11.8 8.8 48.8		Eye Green Hair Black Complexion Tan Teeth Good		60 18 17 17	
175 137 137 67		26.4 11.8 8.8 48.8		Eye Green Hair Black Complexion Tan Teeth Good		60 18 17 17	

DESCRIPTIVE

175 137 137 67	26.4 11.8 8.8 48.8	175 137 137 67		26.4 11.8 8.8 48.8		60 18 17 17	
		175 137 137 67		26.4 11.8 8.8 48.8		60 18 17 17	

175
 137
 137
 67

26.4
 11.8
 8.8
 48.8

60
 18
 17
 17

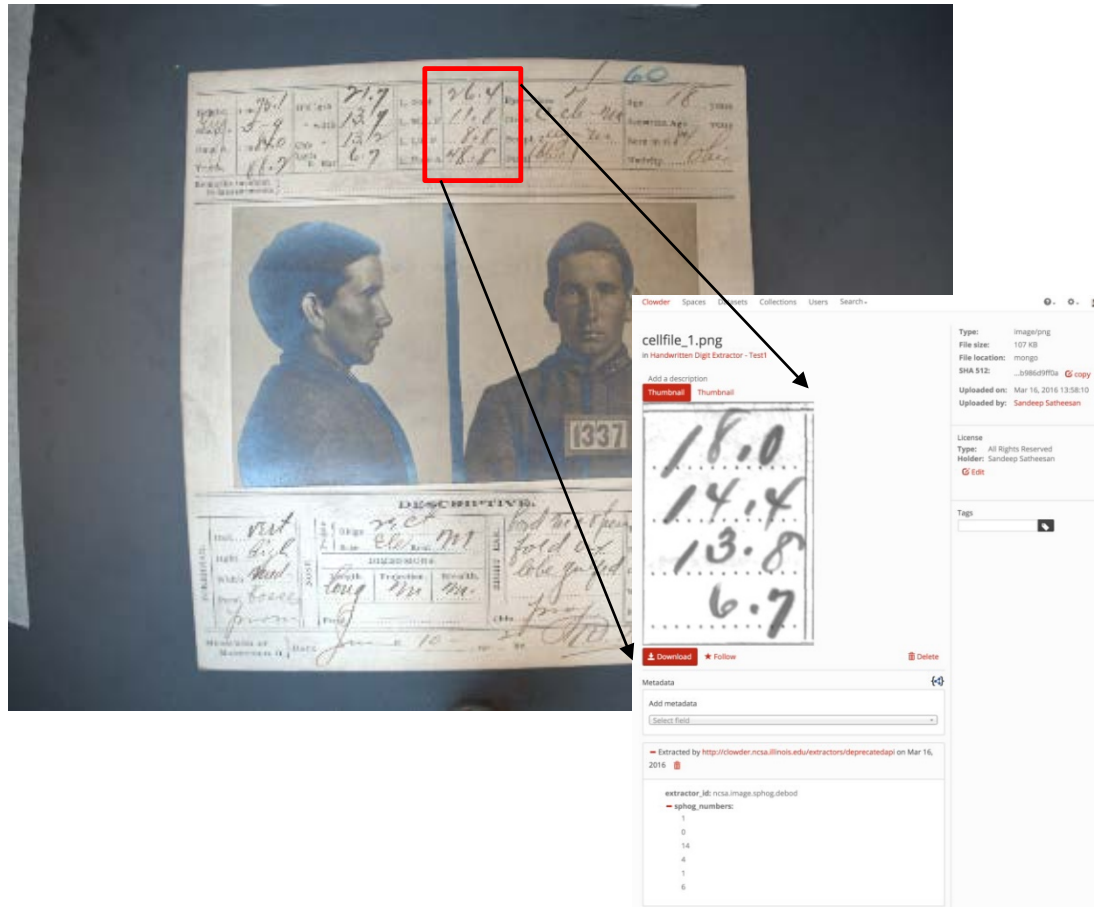
175
 137
 137
 67

26.4
 11.8
 8.8
 48.8

60
 18
 17
 17

OHIO STATE REFORMATORY.	
NAME <i>Wesley Legg</i>	Roll No. <i>1337</i>
ALIAS <i>Pike</i>	Color <i>White</i>
CRIMES <i>Robbery</i>	Age <i>27</i>
COMMITMENT <i>Jan 1895</i>	Discharge <i>Nov 1895</i>
REASON OF REENTRY: <i>REENTRY</i>	
<p>SCARS, MARKS, ETC.</p> <p><i>On neck no scar</i></p> <p><i>On face no scar</i></p> <p><i>On back no scar</i></p> <p><i>On arm no scar</i></p> <p><i>On leg no scar</i></p>	

“A Side”: Segment cells, extract digits, identify digits.
(a well studied problem related to MNIST benchmarks)



“B Side” : OCR on handwritten text
Notoriously difficult – but some constraints:
Descent field has limited entries
One writer for large set of cards

OHIO STATE REFORMATORY.

NAME *Wiley, Legg* REG. NO. *1337*

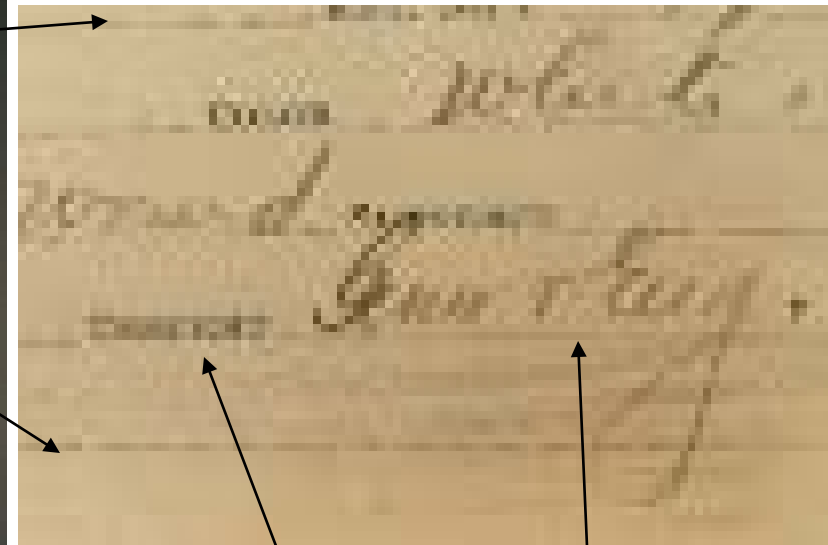
ALIAS *Pike* COLOR *White*

CITY *Cincinnati* STATE *Ohio*

IDENTIFICATION *Yarn band* SIGNATURE *Wiley Legg*

MARKS, SCARS, ETC.

Number	Mark
II	Get not no band 1 ph I at
III	Get not no band 2 ph I at
	Get not no band 3 ph I at



Field word

Field value

Crop grayscale photograph to main card area:
take sum of pixels as profile and apply threshold

INSTITUTION OHIO STATE REFORMATORY. Exp. No. 1131

NAME *Frank Redman*

CRIME *Franklin Burg & Larceny* Race *Indy*

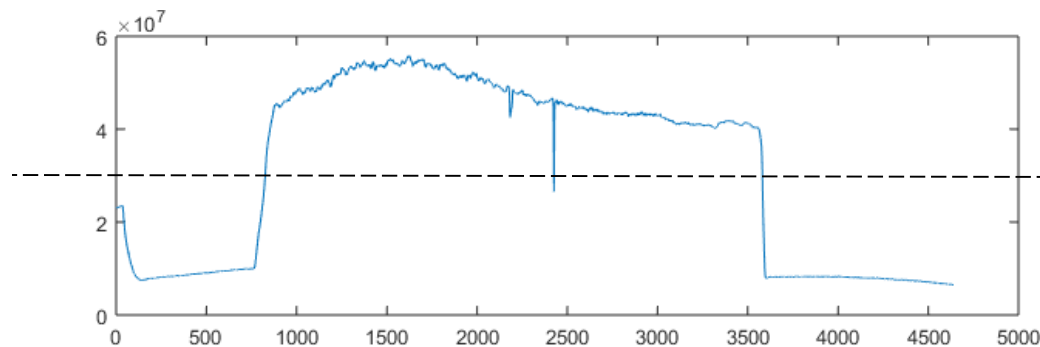
OCCUPATION *laborer* Demeanor *Negro*

REMARKS *Admitted Prison Department & Larceny Oct. 11, 1911.*

MARKS, SCARS AND MOLES

Number	Description	Number	Description
I	Cic. rect 1 & post 4		
	under cut left eye.		
II	Cic. angular 3 ear sup		
	4.5 over wrist at ex.		
IV	Cic. rect 1 thor 7 under		
	beat 12 left of med.		

NOTE.—When filling out this book of Instructions, see note as to correct method of general description, see also particularly as to Marks, Scars, etc. See pages 19 to 21. Use abbreviations as given in instructions.



-take subregion near 'DESCENT'

(hard code expected location)

-binarize

(threshold within smaller local region)

-rotate

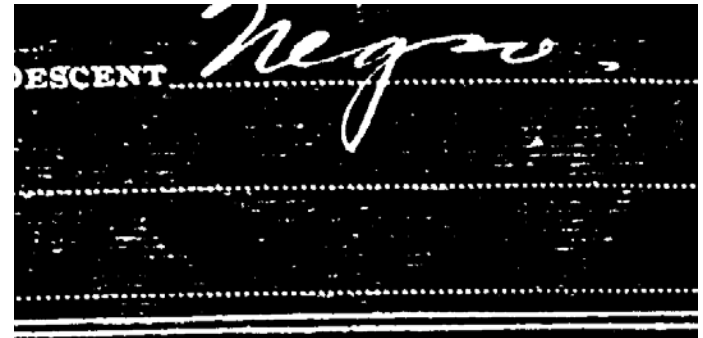
(just a bit to find high/low row sums)

-denoise

(remove small components)

*-find field word by matching profile to
a template (OCR tools didn't work
well)*

*-extract field value by taking window
next to word*



Word spotting

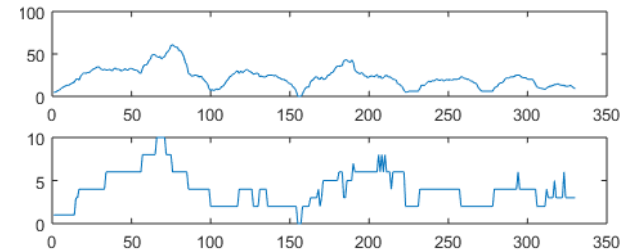


For each test image:

1 Get profiles

Sum of column

*Sum 0-1 or 1-0
transitions in
column*

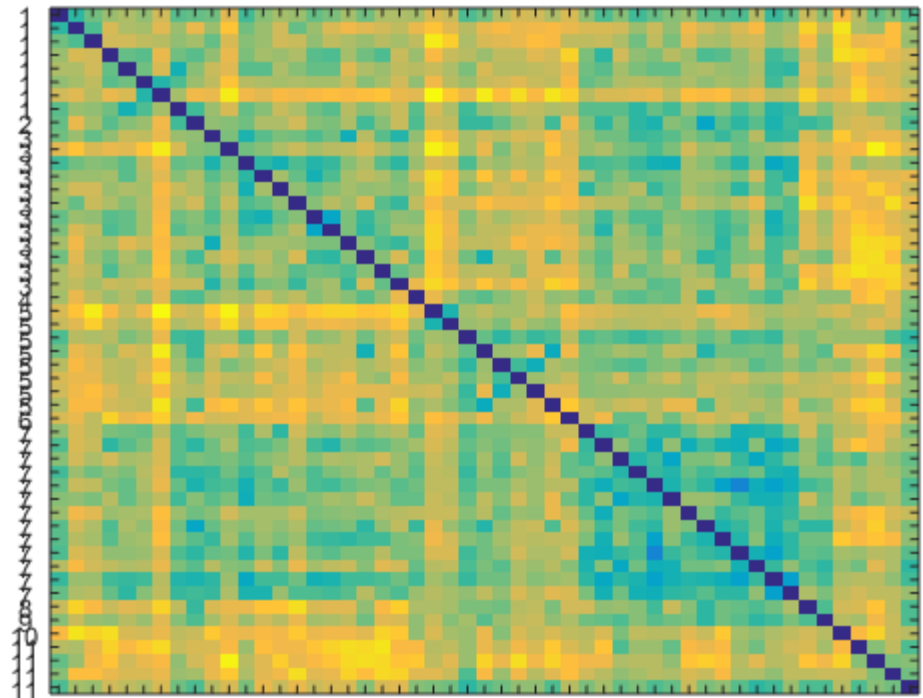


2 Match to known templates: (*gathered 51 templates, 11 nationalities from 1901; testing on 1902*)

2.1. Interpolate to template size

2.2 calculate a distance metric (Euclidean distance is comparable to Dynamic Time warping but much faster)

Dis-similarity matrix of templates (ie distance matrix)



1 1 1 1 1 1 1 2 3 3 3 3 3 3 3 3 3 3 3 3 4 5 5 5 5 5 5 5 6 7 7 7 7 7 7 7 7 7 7 7 8 9 10 11 11 11

English

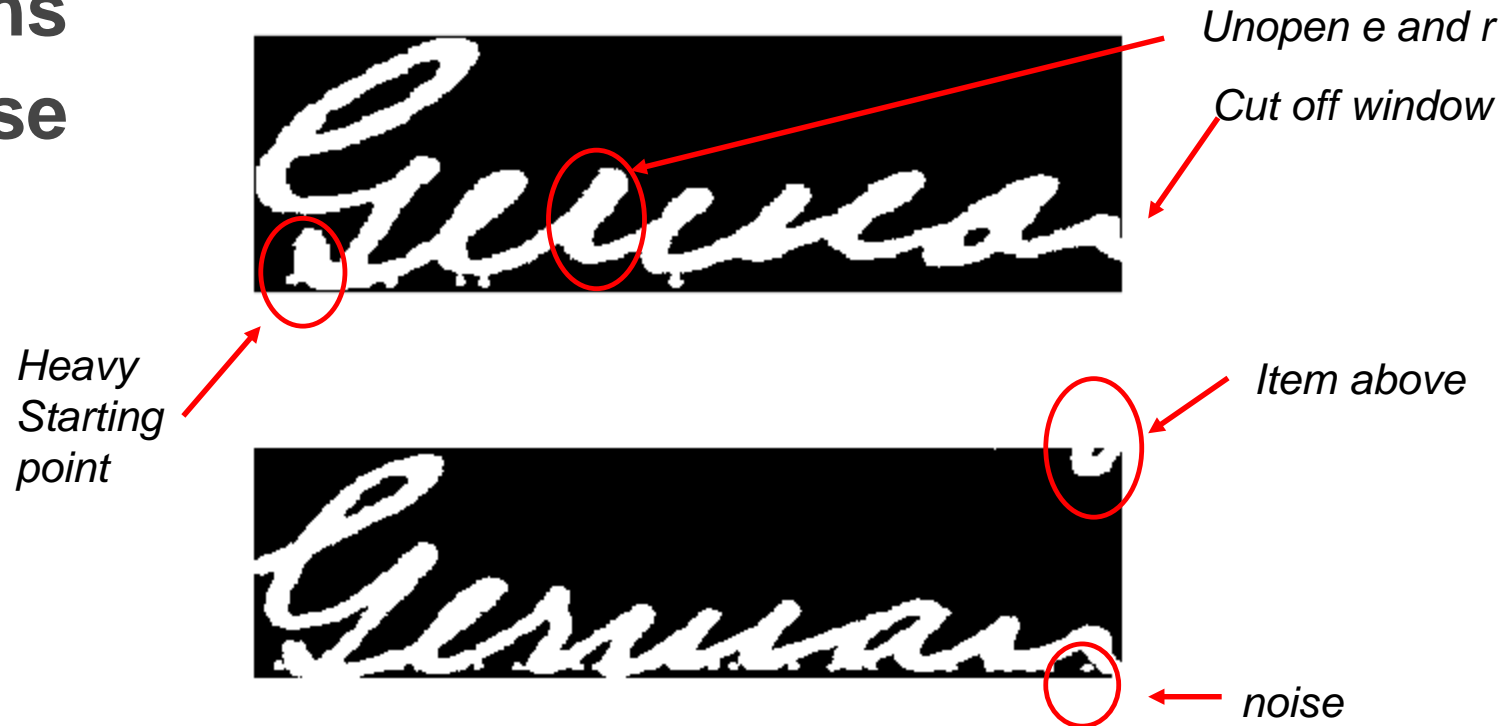
German

Irish

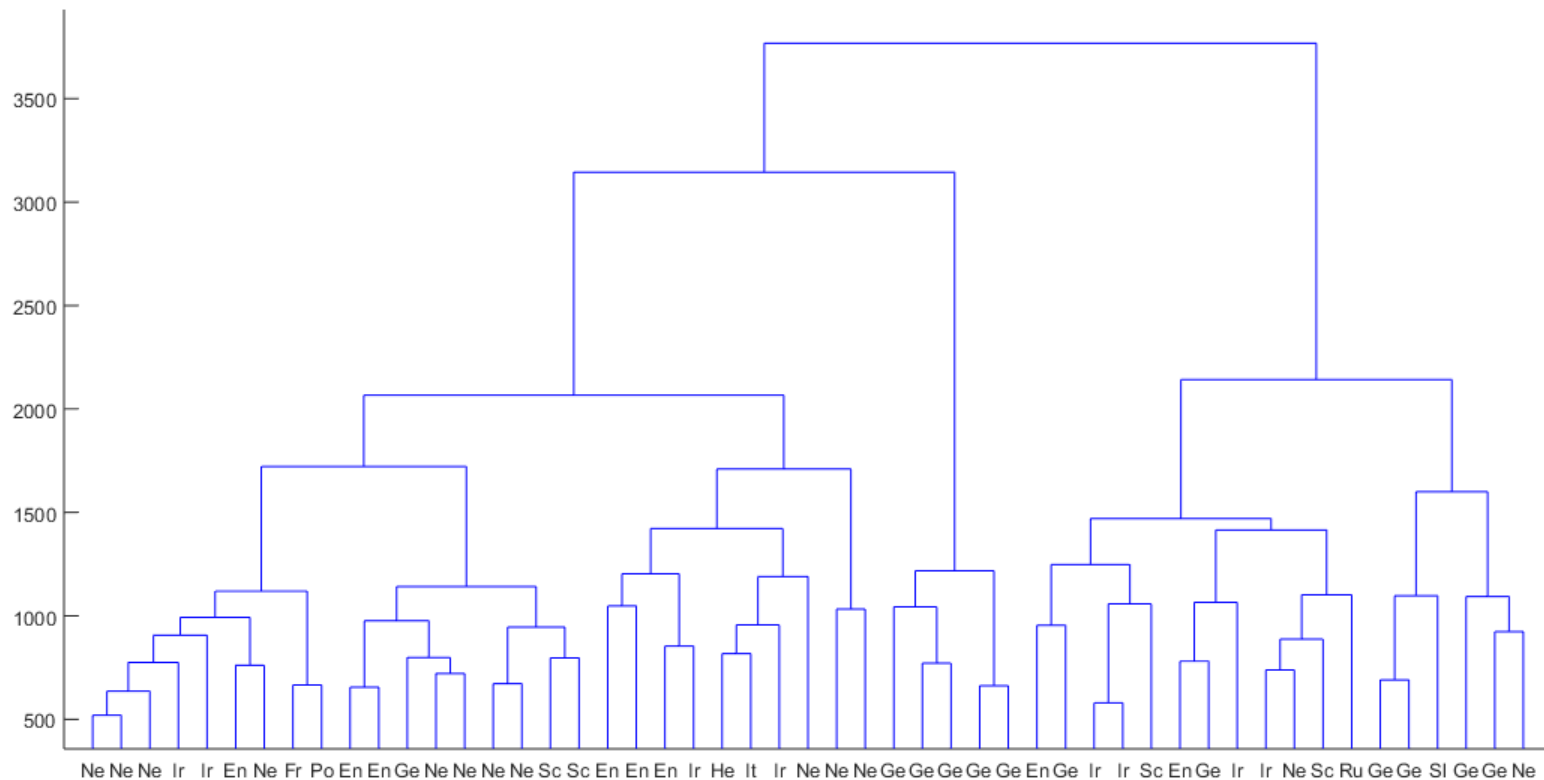
Negro

And a few examples of
'French', 'Hebrew', 'Italian', 'Polish',
'Russian', 'Slavish', 'Scotch'

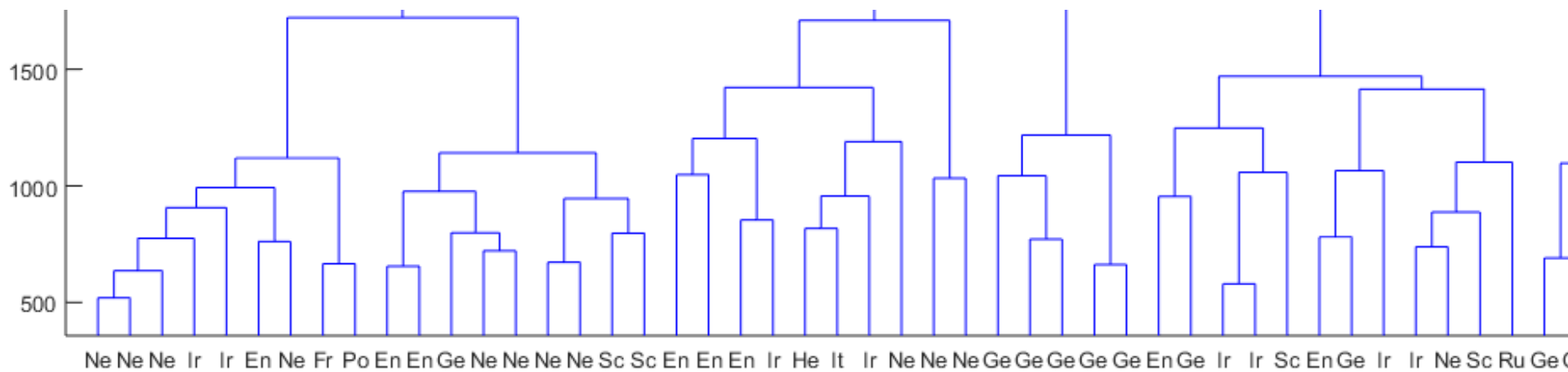
- Variations
- and Noise



- **Template data in hierarchical tree: Ward linkage**



- Zoomed in



Templates: Leave One Out Classifier

- using min distance of out-case to a template
42 out of 51 (so best case estimate of prediction)
- using mean distance of out-case to template set
30 out of 51 (so more likely best case estimate)

Some results with data from 1904:

2097v, Irish 	2098v, Negro 	2099v, Irish 	2101v, Irish 	2102v, Negro 
2103v, German 	2104v, German 	2105v, Italian 	2106v, German 	2107v, Polish 
2108v, Negro 	2109v, English 	2110v, English 	2111v, Negro 	2112v, French 
2113v, Negro 	2114v, Negro 	2115v, Negro 	2116v, Negro 	2117v, Irish 
2118v, German 	2119v, English 	2120v, Scotch 	2121v, Negro 	2122v, English 
2123v, English 	2124v, German 	2125v, Italian 	2126v, Negro 	2127v, German 

Some results with data from
1904:

2097v, Irish



2098v, Negro



2099v, Irish



2101v, Irish



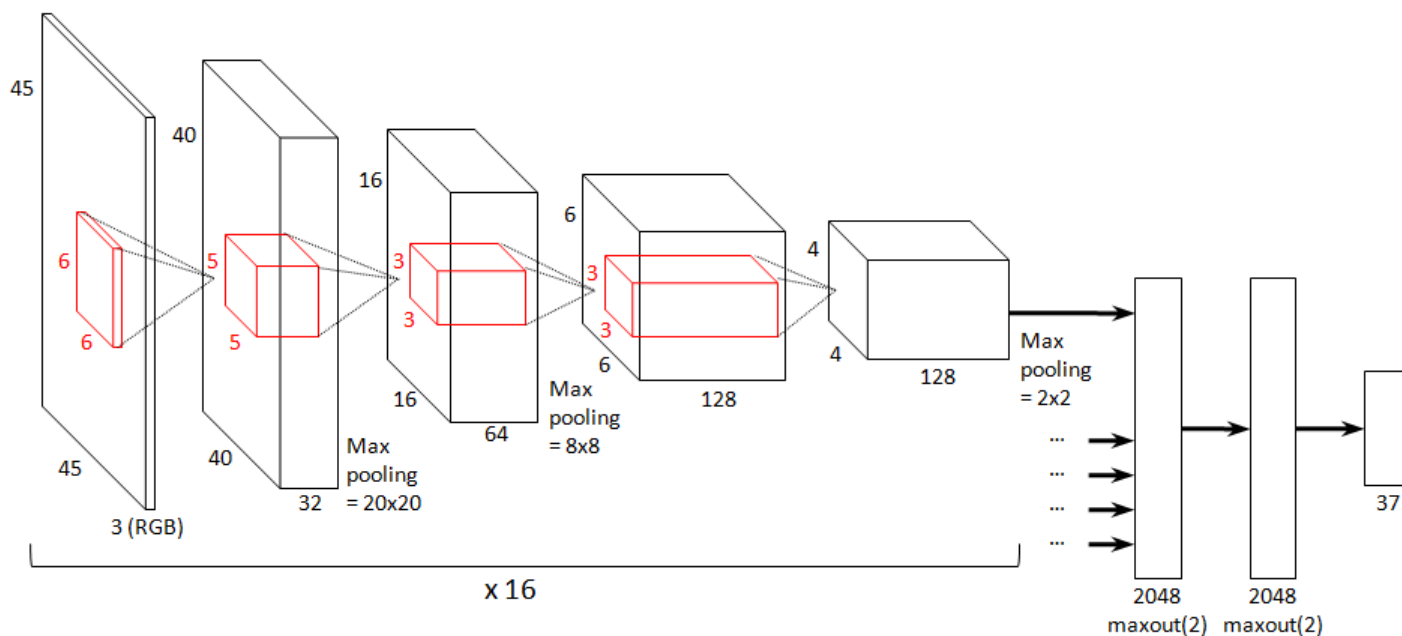
2102v, Negro



*Overall: about 40% correct (for a good sample
set of 30)*

*Errors: 2 images I can't read (eg cut off)
3 image are 2 words (not in templates and
often abbreviated)*

A different feature extractor: convolution network



SciKit python package has a convolution neural network

```
nn2 = Classifier(  
    layers=[  
        Convolution("Rectifier", channels=numch, kernel_shape=(10,10),pool_shape=(2,2)),  
        Convolution("Rectifier", channels=numch, kernel_shape=(6,6),pool_shape=(4,4)),  
        Layer("Sigmoid",units=numalpha2do*4),  
        Layer("Sigmoid",units=numalpha2do*2)  
    ],  
    verbose=False,  
    learning_rate=0.001,valid_set=(Xtrain,Ytrain),  
    n_iter=myiter)  
nn2.fit(Xtrain,Ytrain)
```

Future: use word spotting and convolution network together

