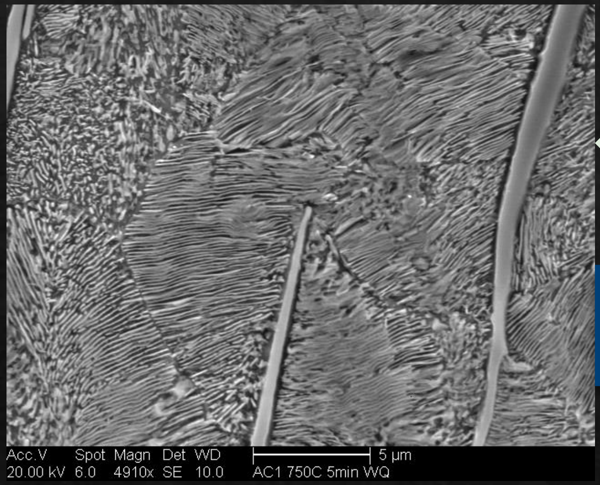
Computer vision and machine learning making the processing-microstructure-property connection

CMU: Nan Gao, Elizabeth A. Holm

NETL: Michael Gao, Youhai Wen, Zongrui Pei

The structure – property connection



Inverse model:
non-unique solution

Forward model:
unique solution

Microstructure
- outcome of processing
- µm scale
- grains, phases, interfaces

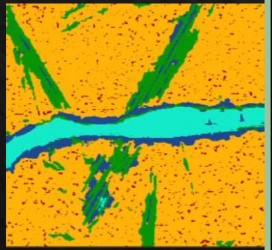
Properties
- outcome of microstructure
- macroscale
- mechanical, chemical, electronic, optical, …

Carnegie Mellon University

Two approaches to the structure-property link

- **Feature-based machine learning:**

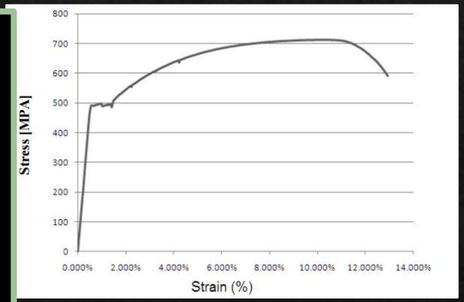
microstructure



features

random forest ML

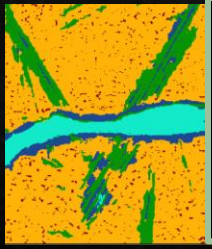
properties



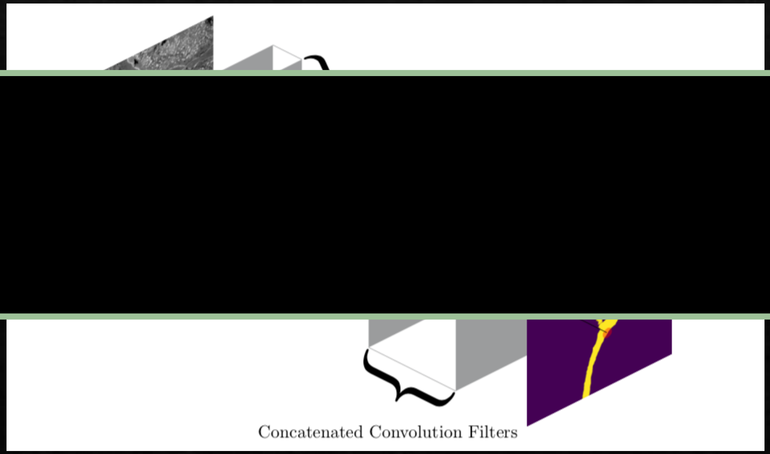
- Interpretable
- Capitalizes on domain knowledge and expertise

- **Holistic machine learning:**

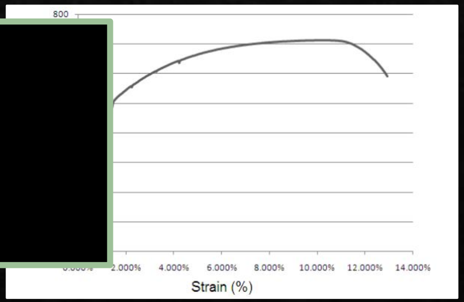
microstructure



convolutional neural network ML



properties



- Powerful
- Black box

Project outline

1. **Data collection**

Archived data
-or-
Newly collected data
-or-
Synthetic data

2. **Image representation**

Constituent segmentation and feature quantification
-or-
CNN feature vector

3. **ML for structure-property connections**

Random forest:
Feature-based ML

CNN:
Holistic ML

Project outline

1. **Data collection**

Archived data
-or-
Newly collected data
-or-
Synthetic data

2. **Image representation**

Constituent segmentation and feature quantification
-or-
CNN feature vector

3. **ML for structure-property connections**

Random forest: Feature-based ML

CNN: Holistic ML

Carnegie Mellon University

Images and metadata

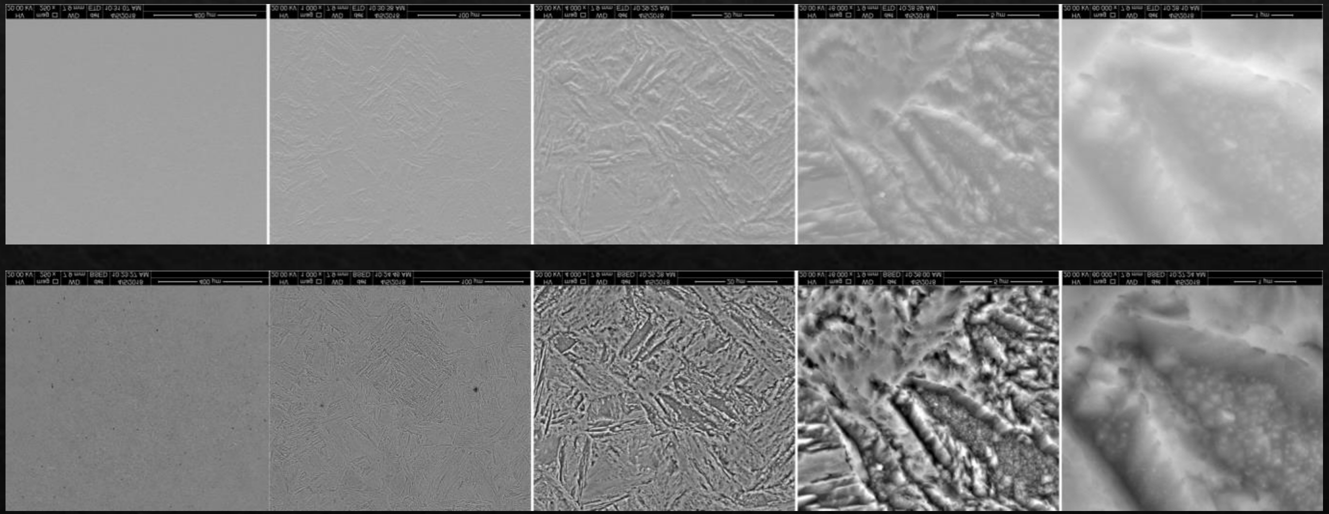
26 9-12% Cr ferritic/martensitic steel alloys

469 SEM images: two detectors, five magnifications

Mechanical test data at multiple temperatures

To create a balanced database, we take 16 images from each class to form a balanced database containing 416 images

A	B
Alloy ID #	# of images
1	16
2	16
3	25
4	16
5	16
6	16
7	17
8	16
9	18
10	17
11	17
12	19
13	21
14	20
15	17
16	18
17	16
18	17
19	17
20	18
94	19
95	28
96	16
97	20
98	16
99	16



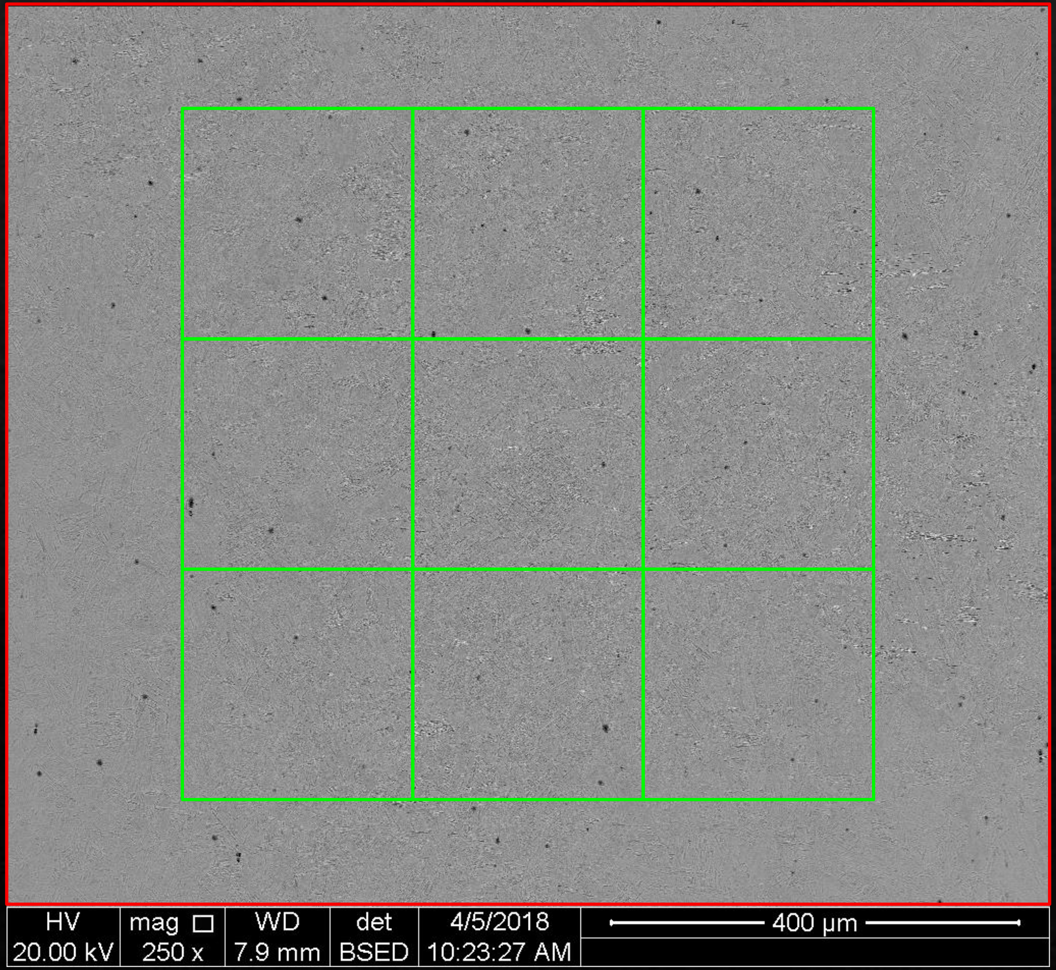
SEM Images

	AH	AI	AJ	AK	AL
1	TT_Temp	TT_YS	TT_UTS	TT_EL	TT_RA
2					
3	24	848	1011	24	61
4	24	836	993	26	66
5	100	801	942	23	67
6	100	800	941	23	63
7	150	785	914	20	65
8	200	769	887	19	64
9	200	768	890	19	61
10	250	759	872	22	68
11	250	758	872	19	59
12	300	744	849	18	60
13	300	743	849	18	62
14	350	731	835	18	67
15	350	730	831	18	65
16	400	708	803	17	61
17	400	712	809	18	67
18	450	680	776	23	69
19	450	686	776	21	63
20	500	642	720	25	76
21	500	635	717	21	67
22	550	567	642	37	84
23	550	568	643	26	71
24	600	494	569	32	81
25	600	474	547	39	77
26	650	396	468	37	83
27	650	390	452	39	85
28					
29					
30					
31					
32					
33					
34					
35					

YS at 24°C

Carnegie Mellon University

Data Processing



Two datasets are created:

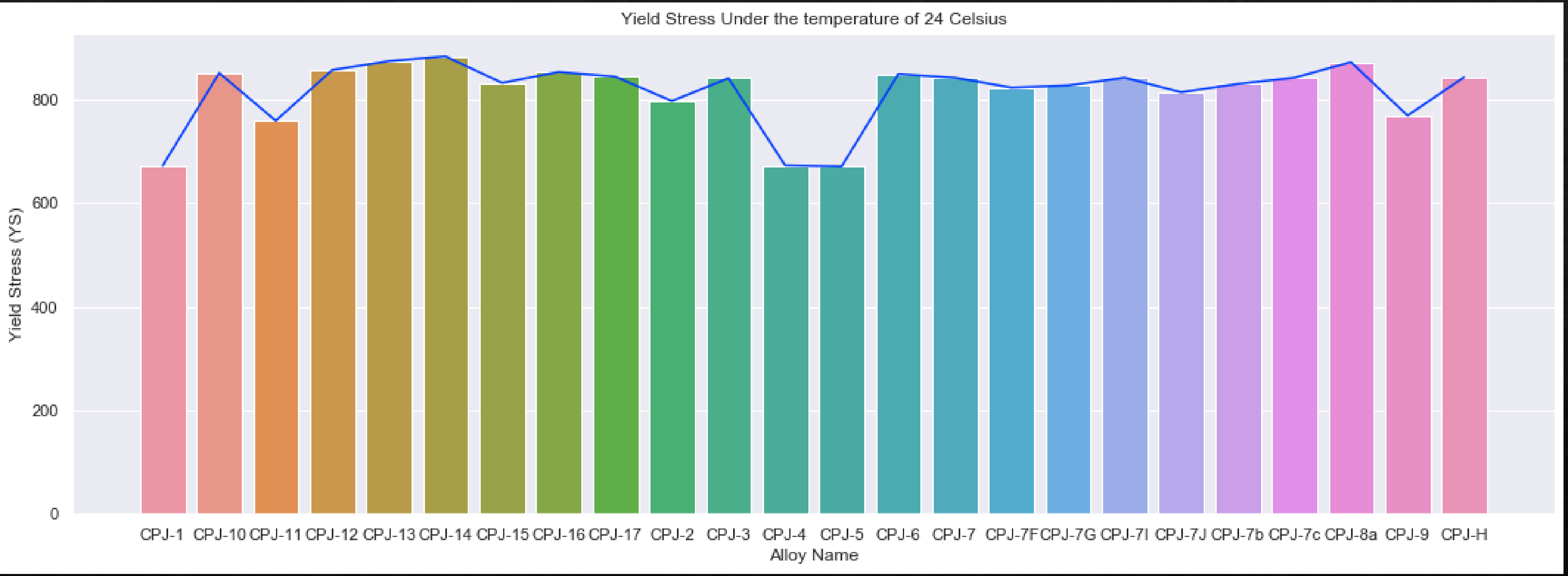
1. NETL-416: 1024 x 943 – full images
2. NETL-3744: 224 x 224 – cropped images

Data augmentation (rotation, scaling, shift) can be used to increase the data volumes as well.

Carnegie Mellon University

Data Preprocessing

	image_path count	unique	...	image_name top	freq
0	20	20	...	CPJ-1	20
1	16	16	...	CPJ-10	16
2	17	17	...	CPJ-11	17
3	17	17	...	CPJ-12	17
4	19	19	...	CPJ-13	19
5	21	21	...	CPJ-14	21
6	20	20	...	CPJ-15	20
7	17	17	...	CPJ-16	17
8	18	18	...	CPJ-17	18
9	28	28	...	CPJ-2	28
10	16	16	...	CPJ-3	16
11	20	20	...	CPJ-4	20
12	16	16	...	CPJ-5	16
13	16	16	...	CPJ-6	16
14	16	16	...	CPJ-7	16
15	16	16	...	CPJ-7F	16
16	17	17	...	CPJ-7G	17
17	17	17	...	CPJ-7I	17
18	18	18	...	CPJ-7J	18
19	16	16	...	CPJ-7b	16
20	25	25	...	CPJ-7c	25
21	16	16	...	CPJ-7d	16
22	16	16	...	CPJ-7e	16
23	16	16	...	CPJ-8a	16
24	18	18	...	CPJ-9	18
25	17	17	...	CPJ-H	17

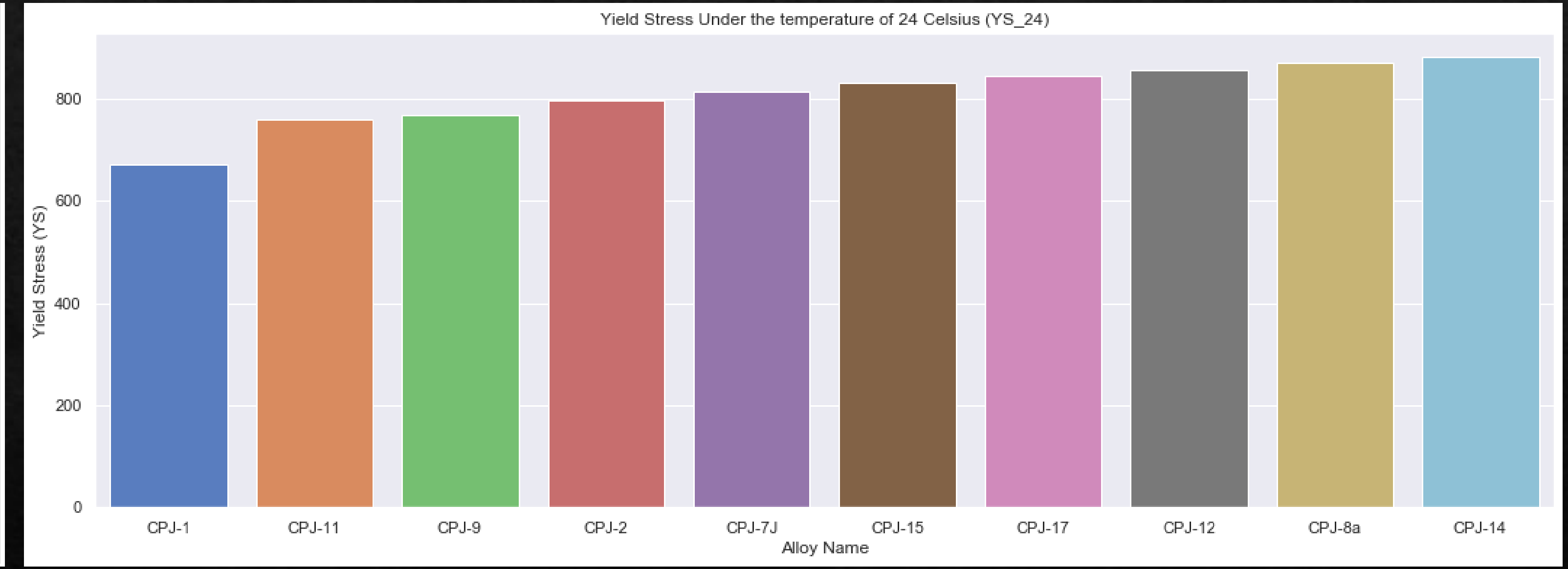


Yield stresses occupy a small range, and some alloys have very similar yield stress

Downselect to 10 alloys

Selection Criterion: At least 10 MPa difference in yield stress

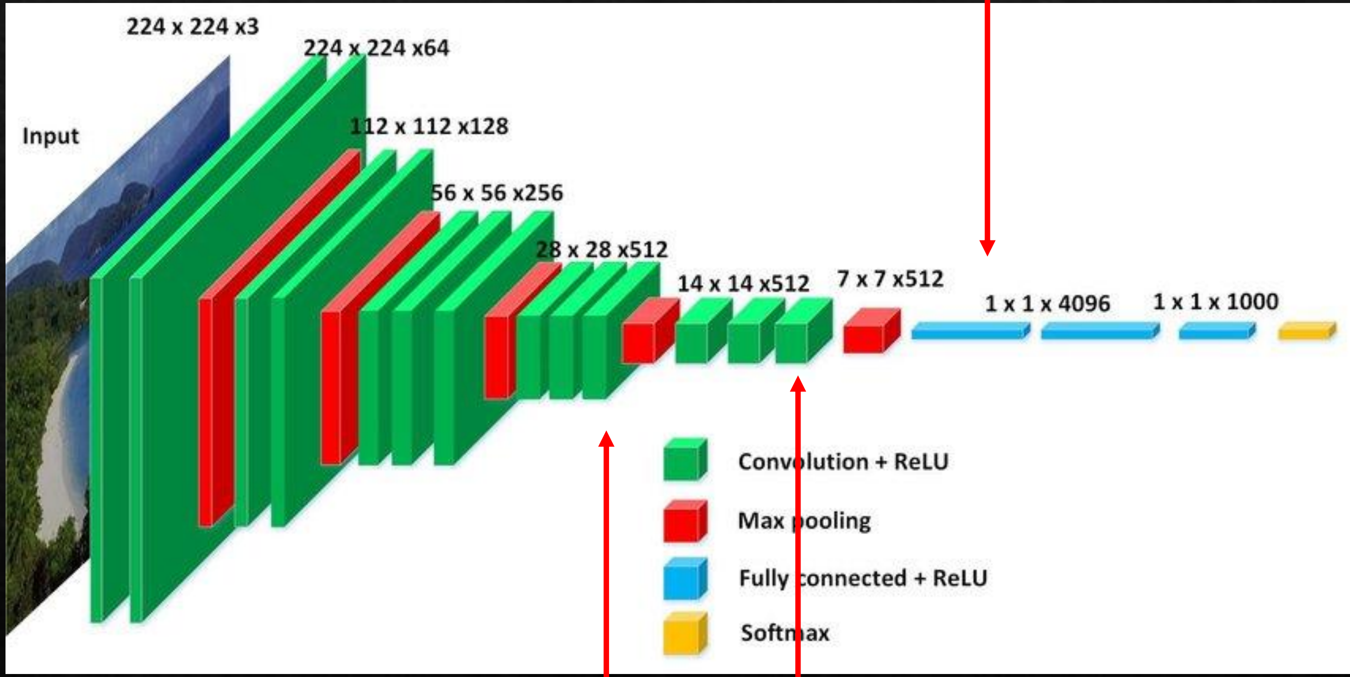
20	94	CPJ–1	672
9	10	CPJ–11	759
8	9	CPJ–9	769
21	95	CPJ–2	797
19	20	CPJ–7J	814
13	14	CPJ–15	832
15	16	CPJ–17	844
10	11	CPJ–12	857
7	8	CPJ–8a	872
12	13	CPJ–14	883



Carnegie Mellon University

VGG 16

Fc1: fully connected layer 1



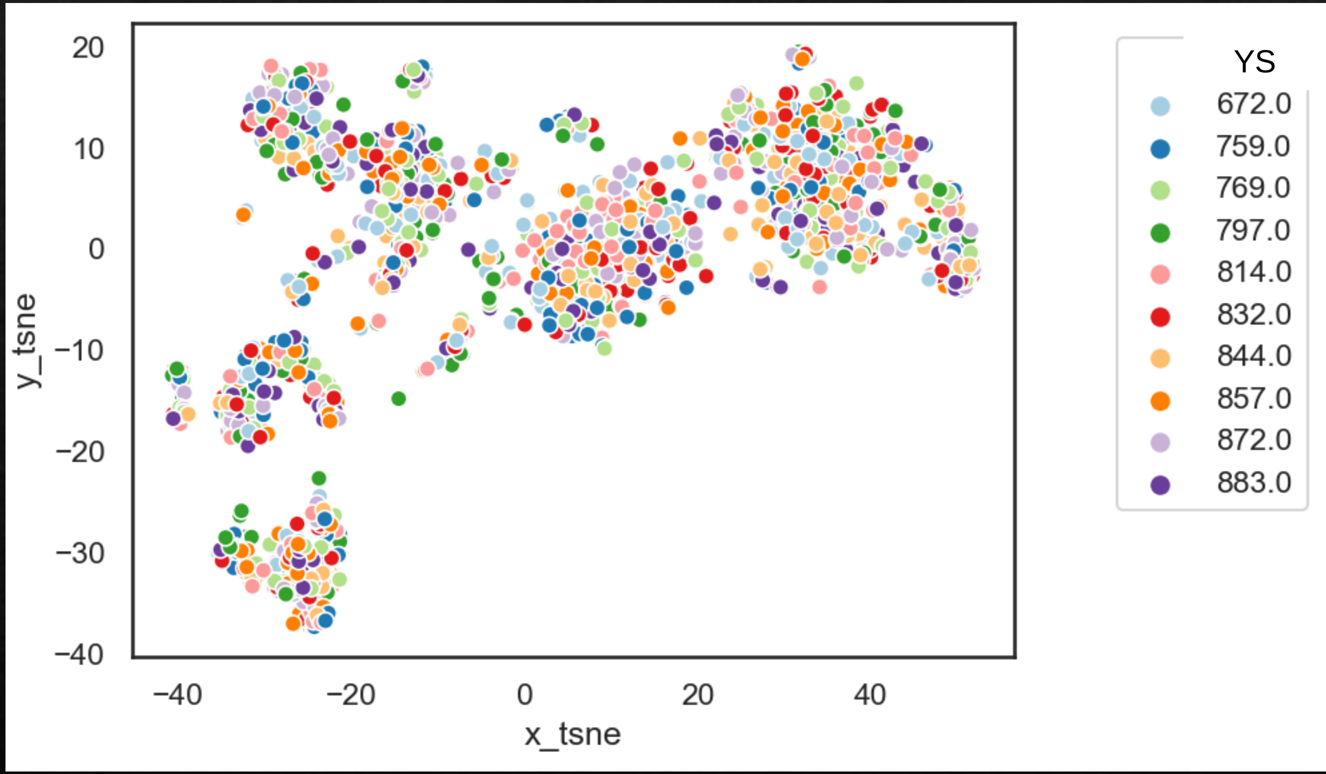
VGG 16 CNN:
Pretrained on the ImageNet database of natural images

b4c3: block 4–conv 3

b5c3: block 5–conv 3

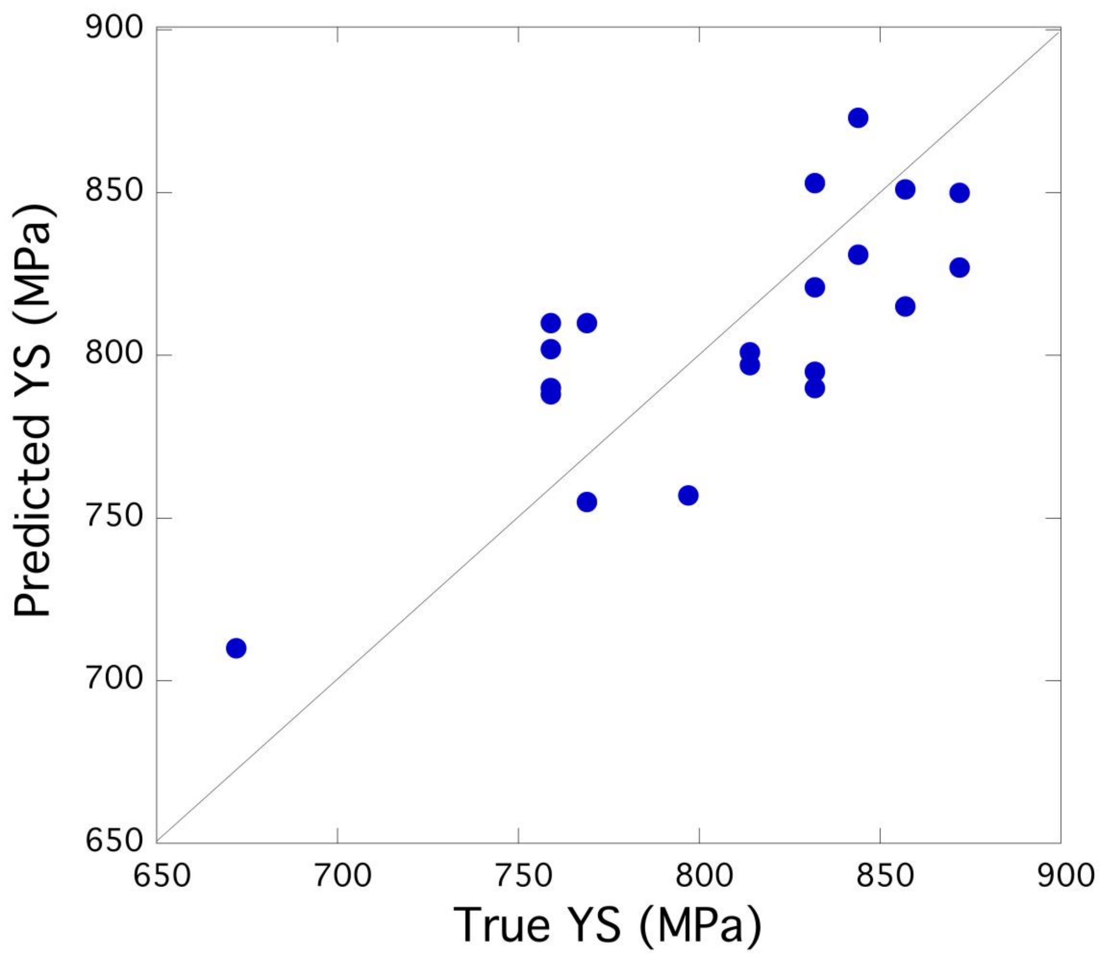
Carnegie Mellon University

Clustering visualization: FC1 features



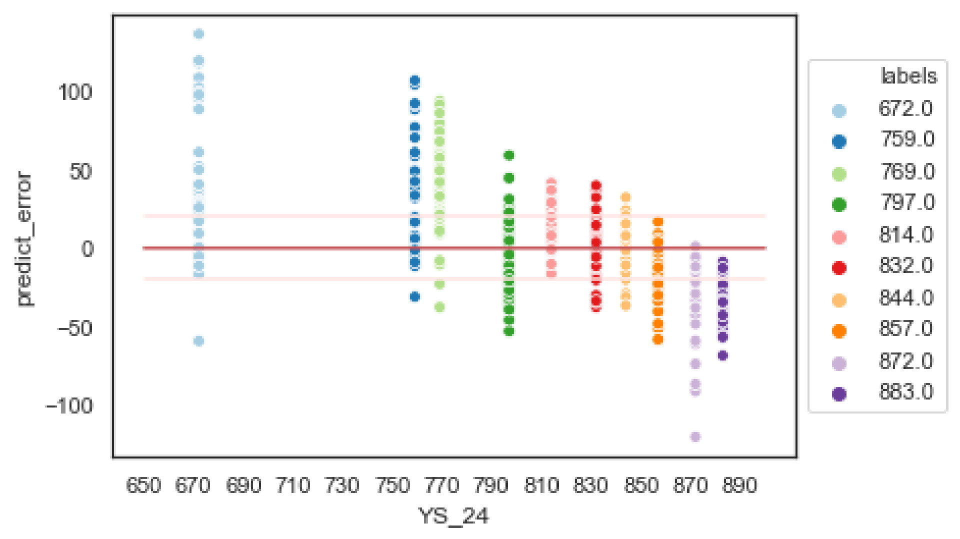
FC1 features are less helpful than the raw image features.

Is there structure we can't see in the CNN features?

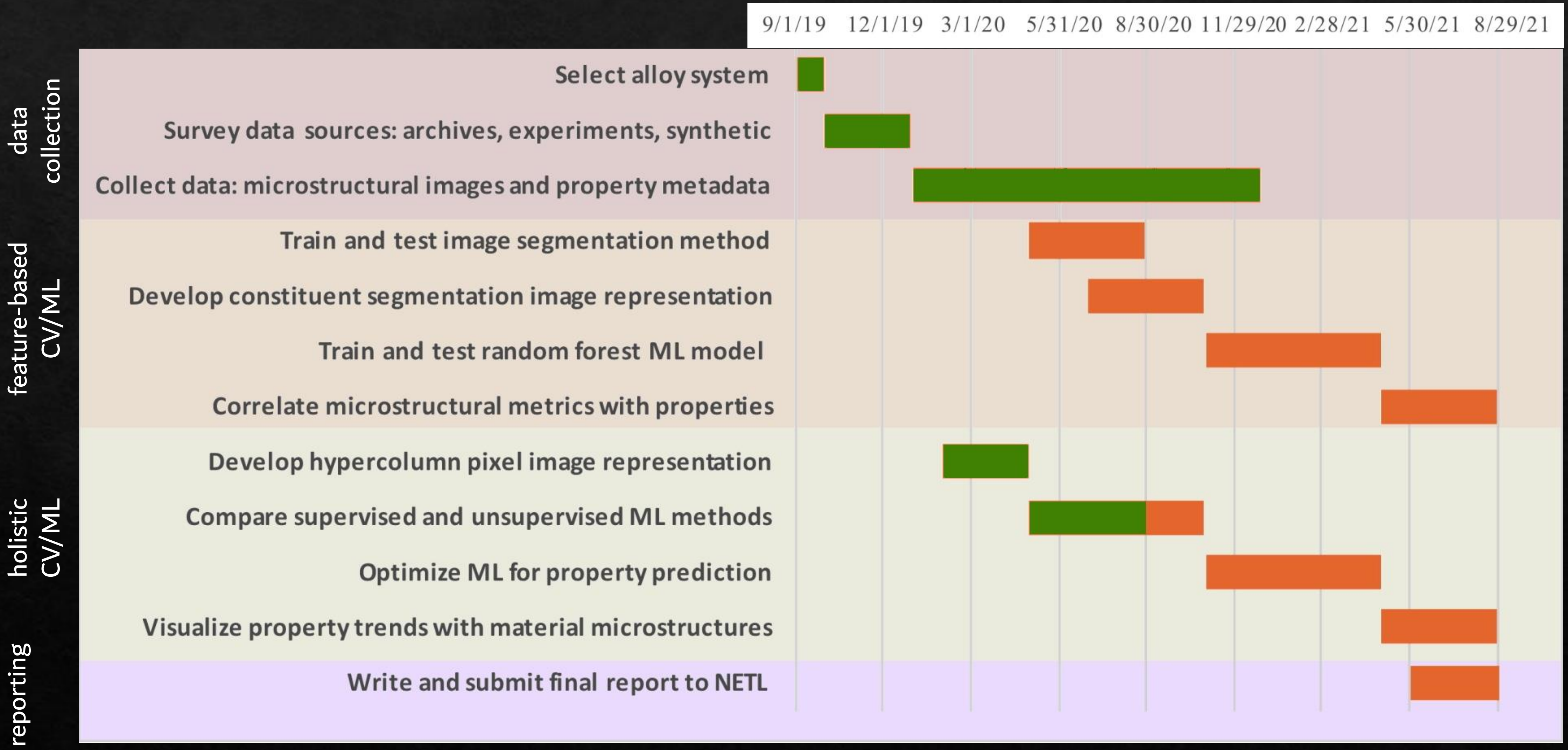


To check, we train a regressor using CNN features.

Promising results – there is a signal!



Project schedule



Carnegie Mellon University

Project goals

1. Collect microstructural image data and property metadata for heat resistant alloy systems
2. Develop material-agnostic CV techniques to extract knowledge from microstructural images.
3. Create ML systems to find relationships between microstructures and property metadata.
4. Analyze and interpret the results to discover new PSP connections.