DRAGNET VISION MODEL

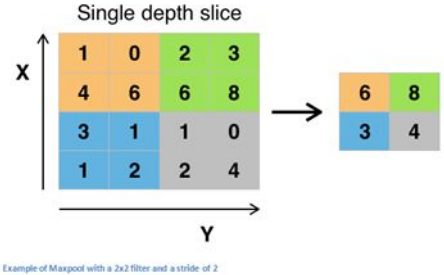
MISSOURI S&T

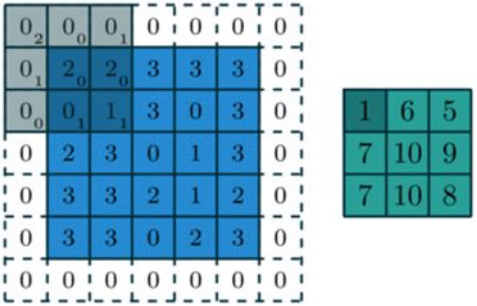
METHODOLOGY

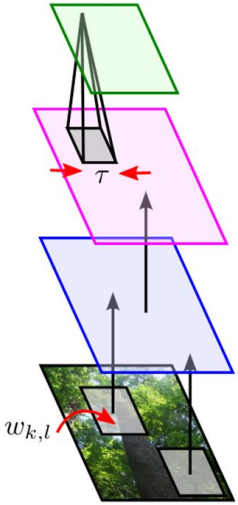
DragNet Vision Model

Convolutional Neural Network Theory

How Do CNNs work?







$$x_{i,j} = \max_{|k|<\tau,|l|<\tau} y_{i-k,j-l}$$

mean or subsample also used — pooling stage

$$y_{i,j} = f(a_{i,j})$$

e.g. $f(a) = [a]_+$

$f(a) = \text{sigmoid}(a)$ — non-linear stage

$$a_{i,j} = \sum_{k,l} w_{k,l} z_{i-k,j-l}$$

only parameters — convolutional stage

$z_{i,j}$ — input image

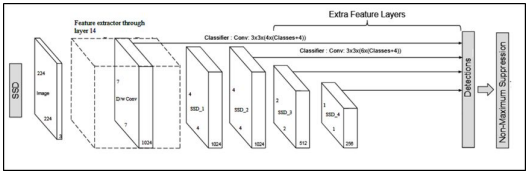
METHODOLOGY

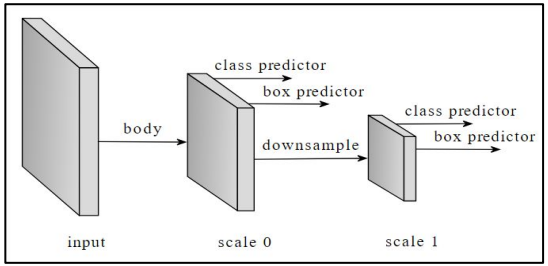
DragNet Vision Model

Model Architecture

Depth	Layer Type	Stride	Filter Shape	Input Size	Parameters
1	Convolution	2	3 x 3 x 32	224 x 224 x 3	864
2	D/w Convolution	1	3 x 3 x 32	112 x 112 x 32	2336
3	D/w Convolution	2	3 x 3 x 64	112 x 112 x 64	8768
4	D/w Convolution	1	3 x 3 x 128	56 x 56 x 128	17536
5	D/w Convolution	2	3 x 3 x 128	56 x 56 x 128	33920
6	D/w Convolution	1	3 x 3 x 256	28 x 28 x 256	67840
7	D/w Convolution	2	3 x 3 x 256	28 x 28 x 256	133376
8	D/w Convolution	1	3 x 3 x 512	14 x 14 x 512	266752
9	D/w Convolution	1	3 x 3 x 512	14 x 14 x 512	266752
10	D/w Convolution	1	3 x 3 x 512	14 x 14 x 512	266752
11	D/w Convolution	1	3 x 3 x 512	14 x 14 x 512	266752
12	D/w Convolution	1	3 x 3 x 512	14 x 14 x 512	266752
13	D/w Convolution	2	3 x 3 x 512	14 x 14 x 512	528896
14	D/w Convolution	2	3 x 3 x 1024	7 x 7 x 1024	1057792
15	Avg Pooling	1	Pool 7 x 7	7 x 7 x 1024	-
16	Fully-Connected	1	1024 x 18	1 x 1 x 1024	-
17	Softmax	1	Classifier	1 x 1 x 18	-
					3185088

Feature Extraction (Howard et al., 2017)





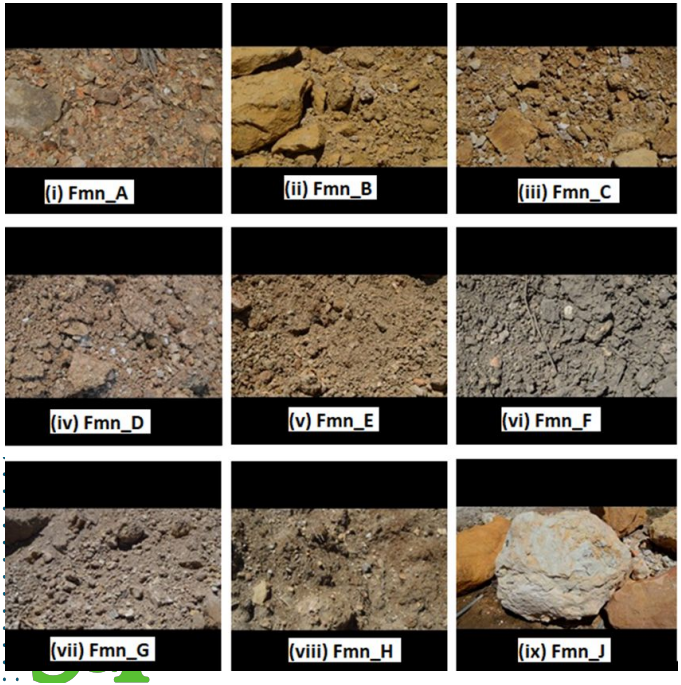
Single Shot Detection (Liu et al. 2016)

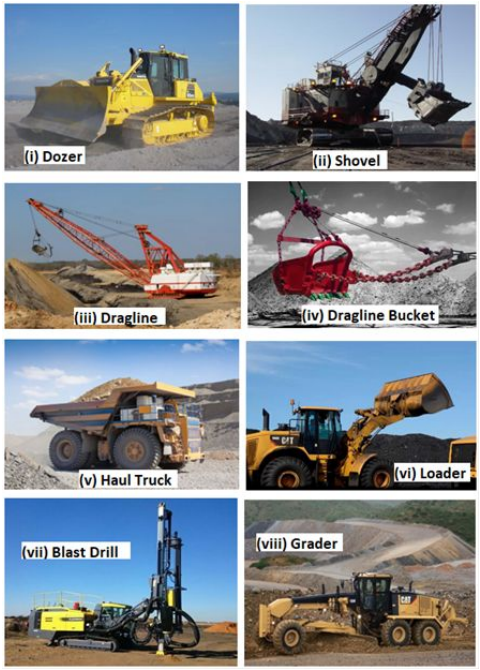
MISSOURI S&T

METHODOLOGY

DragNet Vision Model

Model Input Data

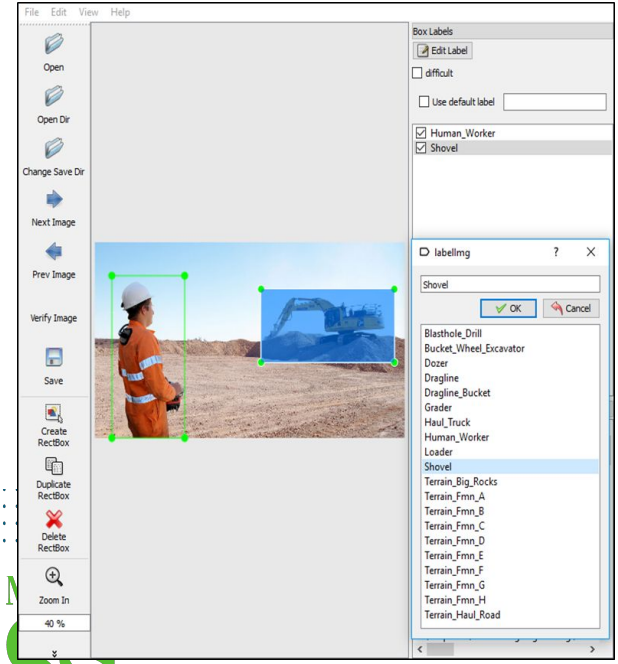


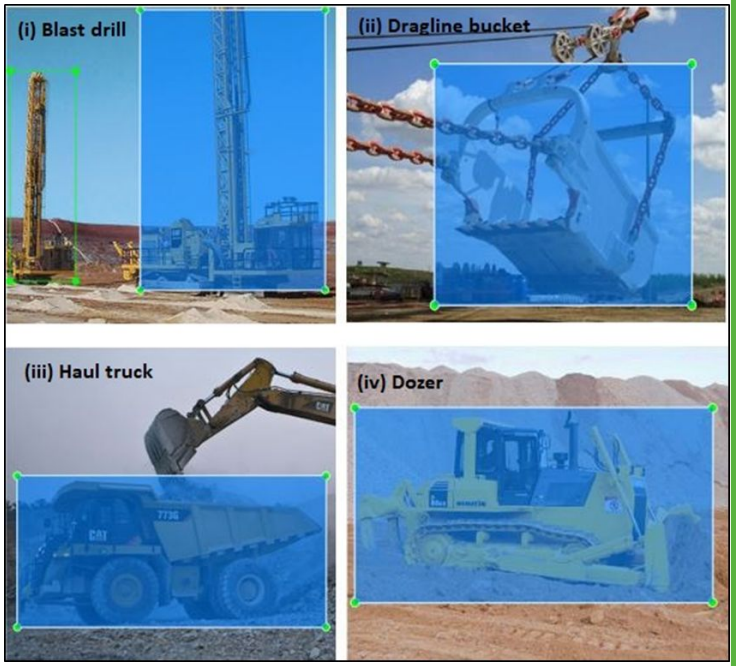


METHODOLOGY

DragNet Vision Model

Image Annotation using LabelImg



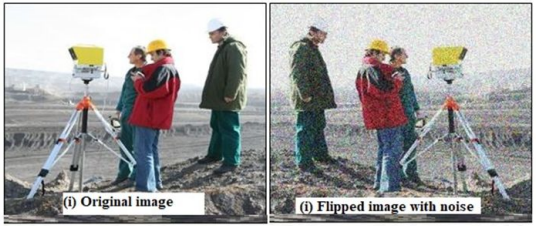


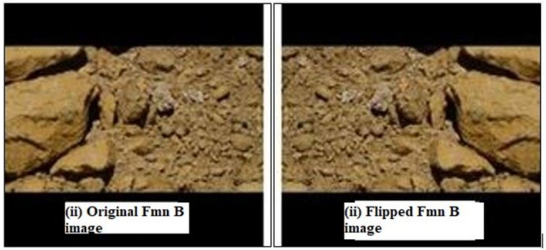
Tzutalin (2017)

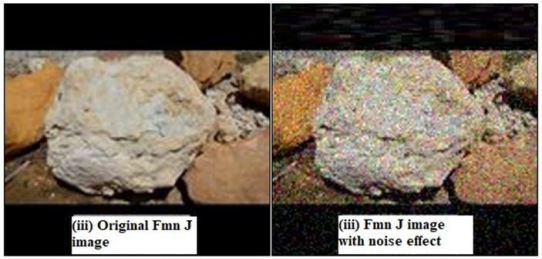
METHODOLOGY

DragNet Vision Model

Data Augmentation

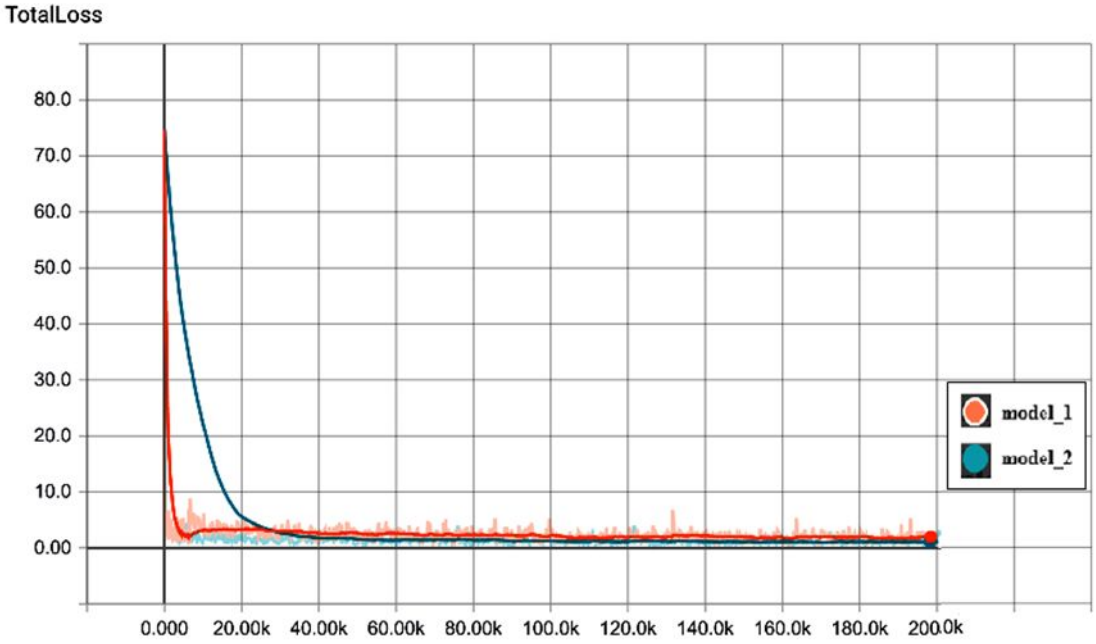






METHODOLOGY

Model Verification



METHODOLOGY

Model Validation

GROUND TRUTH	MODEL PREDICTIONS: Blast Drill	Dozer	Dragline	Dragline Bucket	Fmn_A	Fmn_B	Fmn_C	Fmn_D	Fmn_E	Fmn_F	Fmn_G	Fmn_H	Fmn_J	Grader	Haul Truck	Human	Loader	Shovel	Undetected	Model Accuracy (%)
Shovel	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	22	2	88.0
Loader	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	13	1	1	86.7
Human	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	4	-	-	1	80.0
Haul Truck	-	-	-	-	-	-	-	-	-	-	-	-	-	1	8	-	-	-	2	72.7
Grader	-	-	-	-	-	-	-	-	-	-	-	-	-	13	-	-	-	-	2	86.7
Fmn_J	-	-	-	-	-	-	-	-	-	-	-	-	10	-	-	-	-	-	3	76.9
Fmn_H	-	-	-	-	-	-	-	-	-	-	1	11	3	-	-	-	-	-	4	57.9
Fmn_G	-	-	-	-	-	-	-	-	-	-	20	-	-	-	-	-	-	-	-	100.0
Fmn_F	-	-	-	-	-	-	-	-	-	19		-	-	-	-	-	-	1	-	95.0
Fmn_E	-	-	-	-	-	-	-	-	10	-	5	-	-	-	-	-	-	1	-	62.5
Fmn_D	-	-	-	-	-	-	-	12	-	-	-	-	2	-	-	-	-	-	2	75.0
Fmn_C	-	-	-	-	-	-	19	1	-	-	-	-	-	-	-	-	-	-	2	86.4
Fmn_B	-	-	-	-	-	8	-	-	-	-	-	-	-	-	-	-	-	1	2	72.7
Fmn_A	-	-	-	-	23	-	-	-	-	-	-	-	4	-	-	-	-	-	-	85.2
Dragline Bucket	-	-	-	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1	88.9
Dragline	-	-	8	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	3	72.7
Dozer	-	17	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100.0
Blast Drill	16	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	100.0
																		Average:		82.6

MISSOURI S&T

RESULTS & DISCUSSION

MISSOURI S&T

Results & Discussions

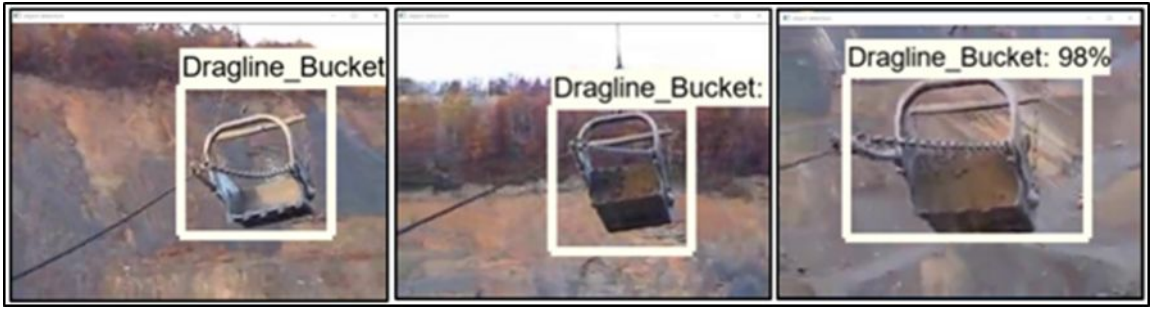
DragNet Application: Bucket Pose Estimation

Some loading phase detection results.

Some hoisting phase detection results.

Results & Discussions

DragNet Application: Bucket Pose Estimation

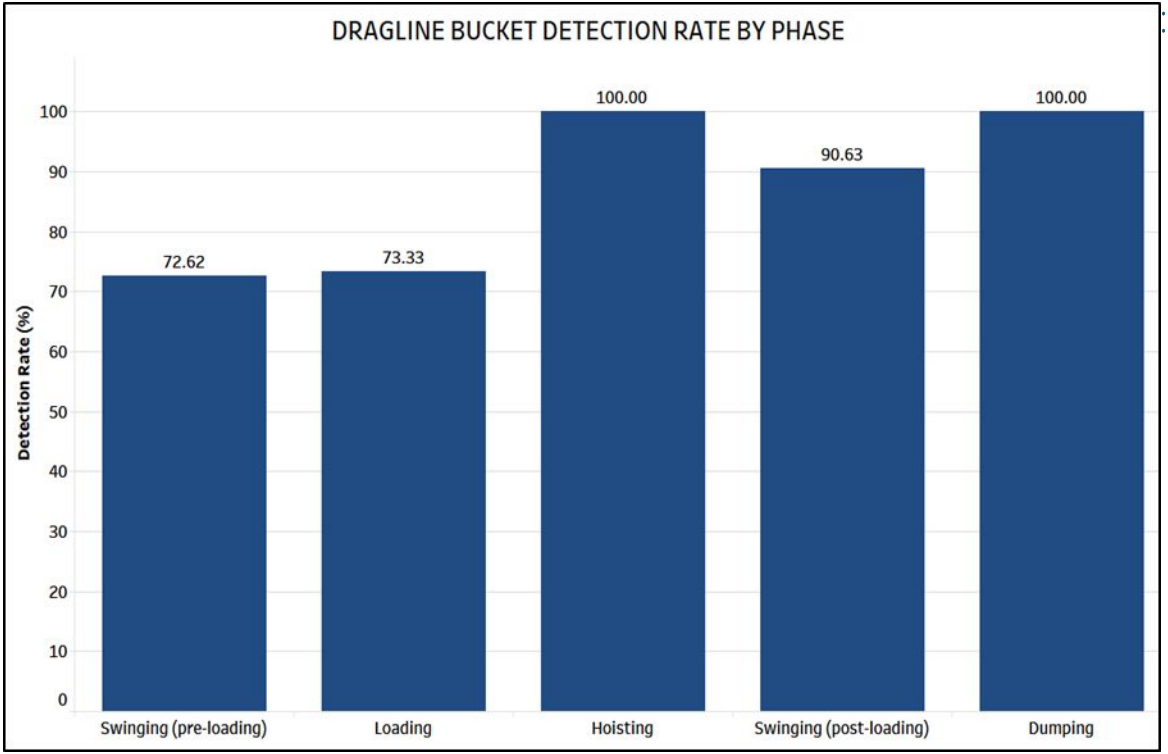


Some swinging phase detection results.

Some dumping phase detection results.

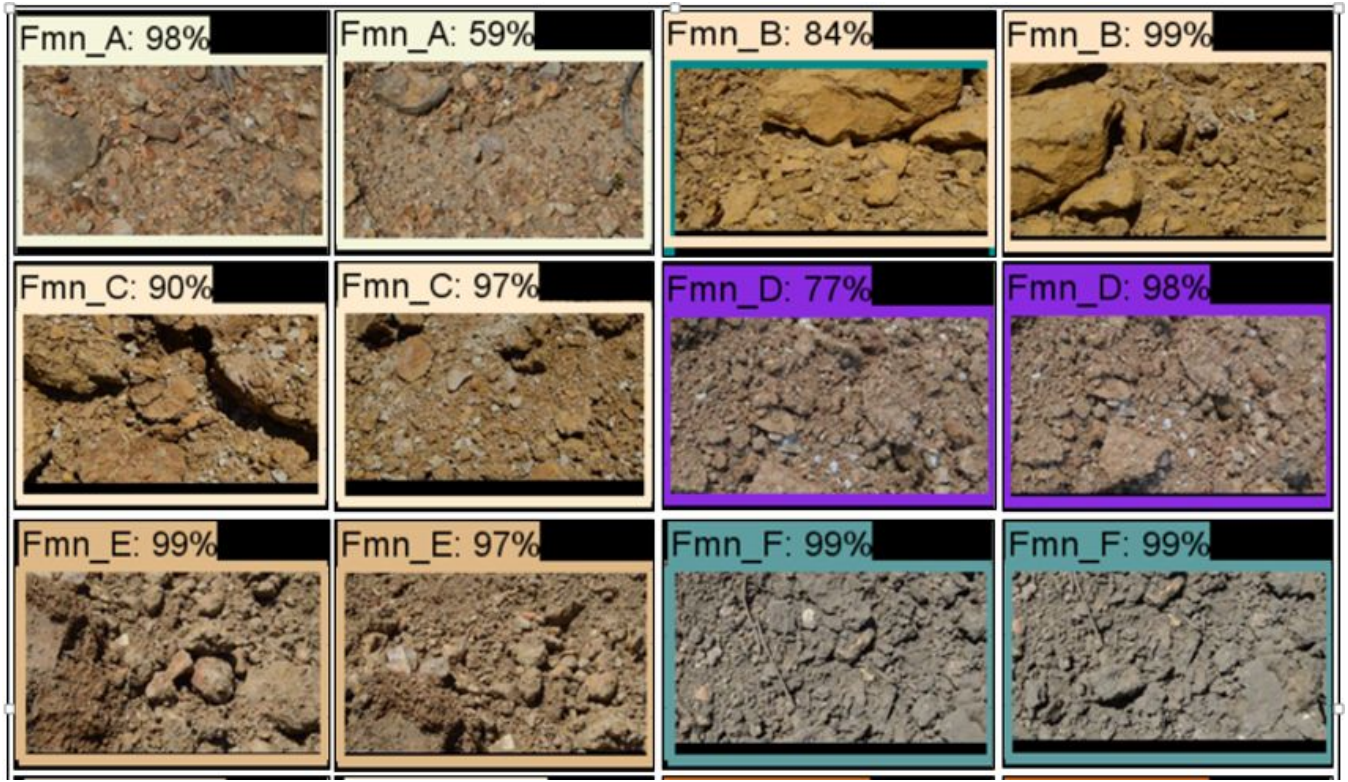
Results & Discussions

DragNet Application: Bucket Pose Estimation



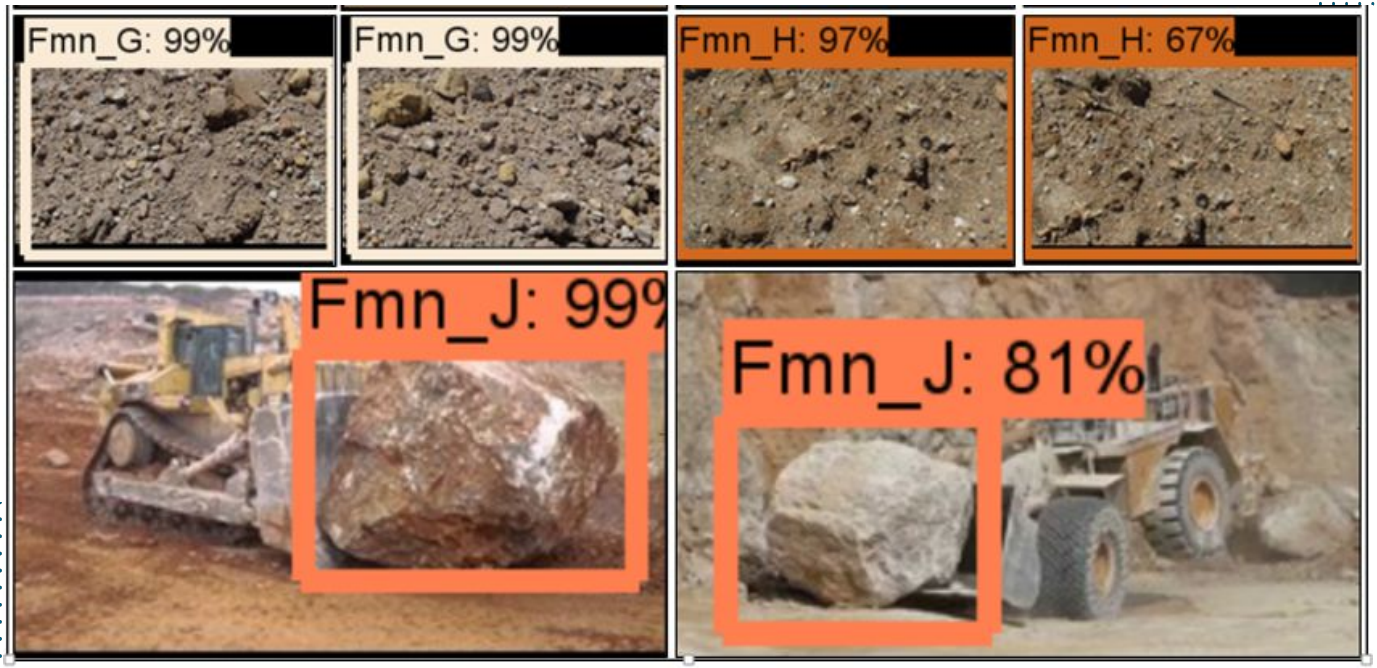
Results & Discussions

DragNet Application: Terrain Recognition



Results & Discussions

DragNet Application: Terrain Recognition



Results & Discussions

DragNet Application: Terrain Recognition

Conclusions – Dragline Vision

1. The DragNet model is able to achieve an 87.32% average detection rate across all operation phases on bucket pose estimation tasks.

2. The DragNet model is able to achieve 80.9% precision and 91.3% recall performance on terrain recognition tasks.

3. While the DragNet model performs considerably well, future improvements will be required to meet minimum performance thresholds for safe operation.

MISSOURI S&T