An Unsupervised Augmentation Framework for Deep Learning based Geospatial Object Detection

By

Devansh Sheth
Ganesh Krishnan Sivaram

Outline

- Introduction
- UA framework
- ROV based approach
- Context based approach
- Experimental evaluation
- Future works

What are we trying to achieve ?

- Aim is to detect geospatial objects with minimum bounding rectangles from remote sensing datasets.
- This can be useful for societal applications like urban planning, census, sustainable development, security surveillance etc.

(a) Input image (b) MBR output

What are the challenges ?

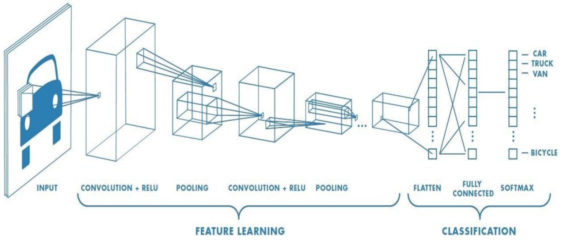
- Detection is challenging because their orientation is heavily mixed and not parallel to orthogonal directions due to topography, planning.

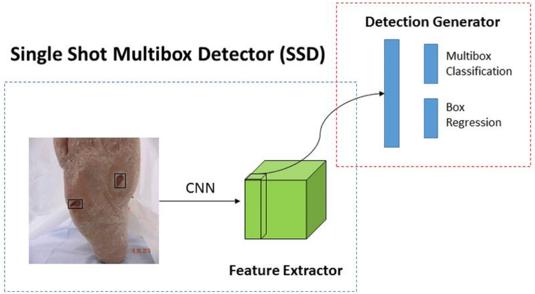
- Also limited training data is available with angle information for most of the objects.

Recent developments

Deep learning frameworks for image processing like :

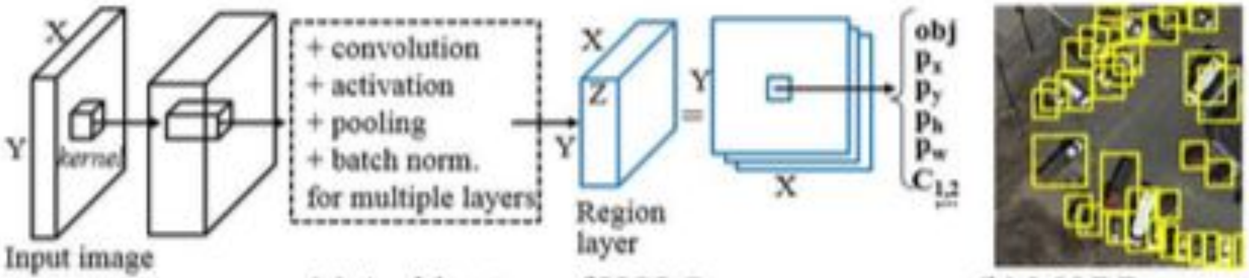
- YOLO (You Only Look Once)
- SSD (Shot Multibox Detector)





YOLO

YOLO gives Minimum Orthogonal Bounding Rectangles which is smallest rectangle having sides parallel to sides of image and cover the object.



(a) Architecture of YOLO

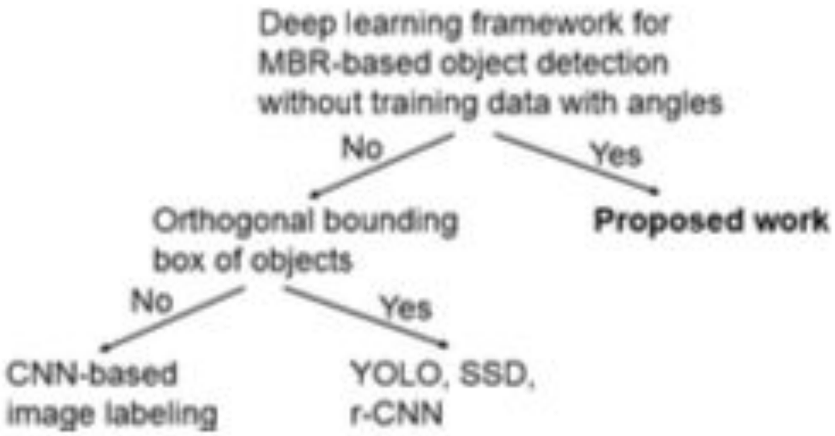
(b) MOBR output

Limitations of deep learning frameworks

- They are not flexible in directions
- Limited availability of training data with angle attributes
- Difficult to generalize it for large scale applications

How to improvise the above frameworks ?

By including angles !!



UA Framework

Authors propose an Unsupervised Augmentation framework for detecting general MBR of geospatial objects.

It tries to solve the problem of unavailability of training data with angles. The 2 schemes presented are:

1. Rotation Vector based scheme
2. Context based theme.

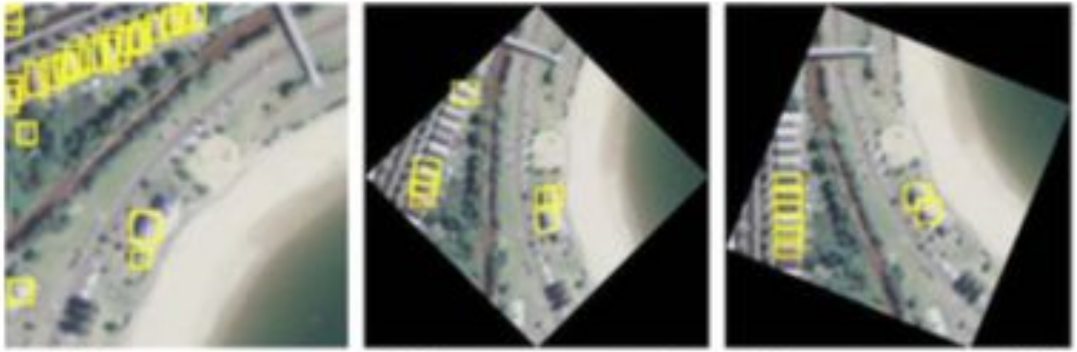
ROV based approach

Rotation Vector based approach augments each test image by rotating it at different angles.

The idea is to get rotated presentations for each object in augmented test data and then use the detected MOBR at different angles to derive best angle and the size of its MBR.

ROV based approach

Now rotating the images generate empty areas around the boundary, which may cause errors in detection. So, for addressing this, 2 completion algorithms are proposed: ROV Reflection and ROV Spatial



(a) No empty area (b) Empty areas at $\pi/4$ (c) Empty areas at $5\pi/16$

ROV Reflection

ROV Reflection completes empty areas by reflecting scenes in the image using the image borders as mirror lines. So, the black space on top left corner is completed using reflections from inside the image with blue line as the mirror line.

ROV Spatial

ROV Reflection still generates non natural things which can potentially affect the final results. To solve this, ROV spatial is used which fills the empty areas using the original scenes. This requires access to large source imagery which was used to generate the test images. Doing this guarantees that each block is filled with actual missing data.

Filtering, projection and grouping

The objects in test image can have different angles, so it is not proper to use detections from single rotation to represent MBRs of all objects in image.

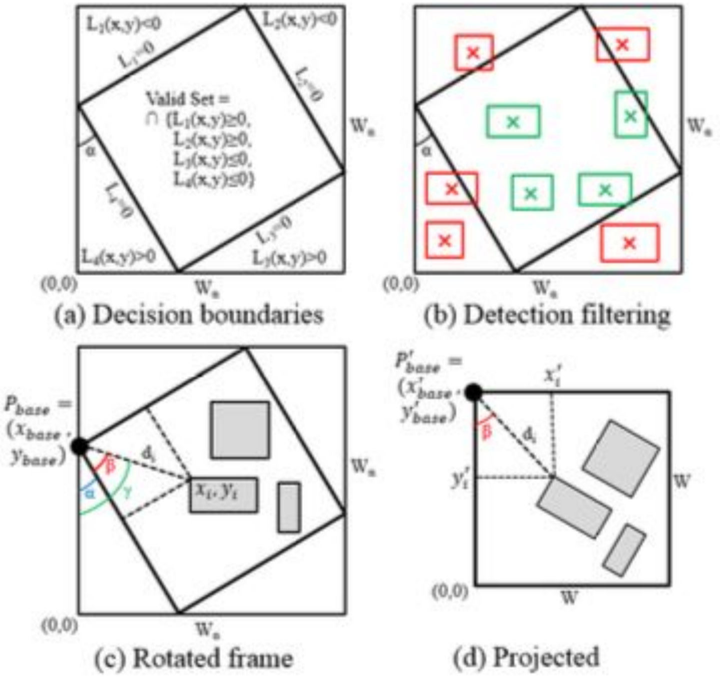
This is where FPG comes into the picture. FPG stands for Filtering, Projection and Grouping.

Filtering, projection and grouping

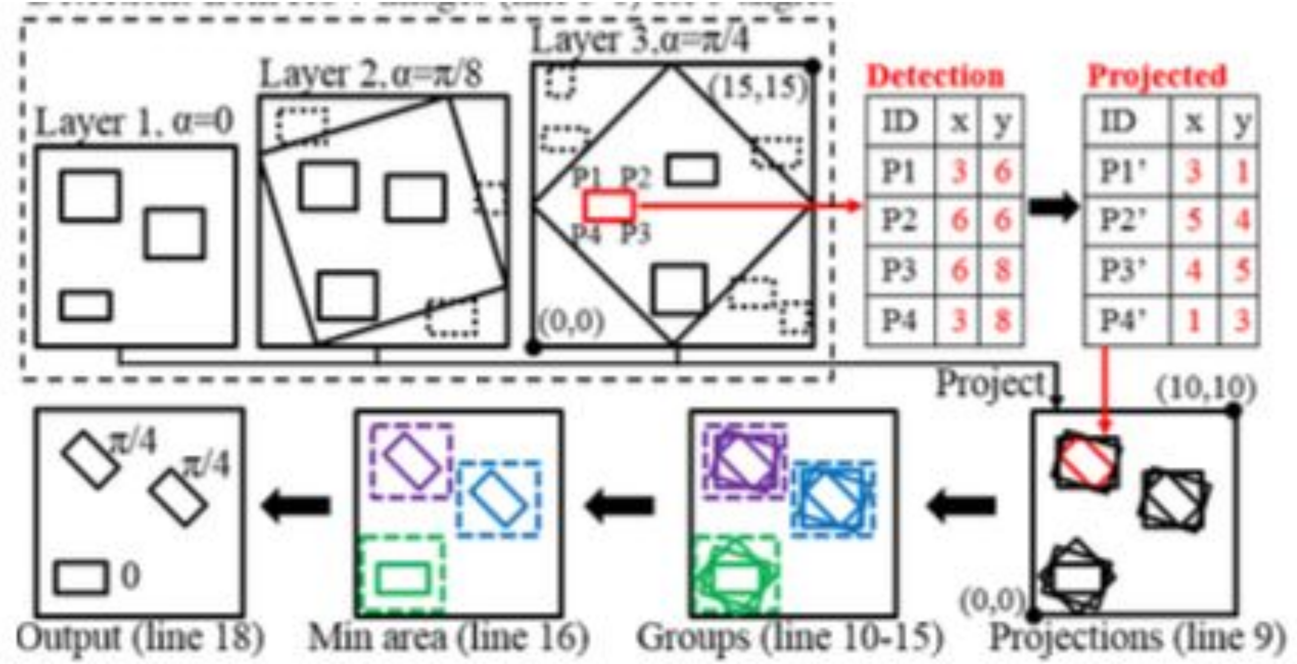
Filtering: Removes detections whose centers are not within the original image.

Projection: This phase maps all detections from rotated image back to original image.

Grouping: This phase clusters detections into groups, each containing detections of same object.



Example



When all 3 phases are completed, we can select the MBR by selecting detection with minimum area in each group.

Context-based Approach

Context based Approach

- No need to find the rotation angle
- Uses spatial context information of certain objects, that can aid in inferring the rotation angles
- Example :

  Road and topographic models serve as context for buildings and farm fields.

Context based Approach

- Building (marked in yellow) are built in context of roads, even if roads are curved.

Context based Approach

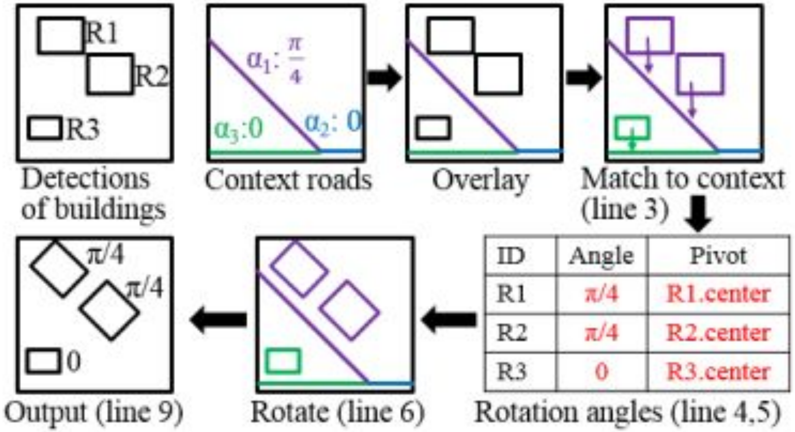
Algorithm 2: Context-based method

Require:

(1) A test image img;
(2) A list of context objects L_{ctxt};
(3) A context-object angle β;

1: L_{result} = **new** List()
2: L_{mobr} = deepCNN(img)
3: **for** $MOBR$ in L_{mobr} **do**
4: c = getContextItem($MOBR$, L_{mobr})
5: α = getContextAngle(c)
6: α' = getRotationAngle(α, β)
7: $MOBR'$ = rotate(obj: $MOBR$, pivot: $MOBR.center$, angle: α')
8: L_{result}.append($MOBR'$)
9: **end for**
10: **return** L_{result}

- β is the radian of the angle between the direction of the context and the object.



Data requirement for UA framework

Table 1: Data requirements for the proposed methods

Requirements	Context	ROV-Reflection	ROV-Spatial
Access to context data	✓		
Access to source test data			✓
Access to plain test data	✓	✓	✓

Experimental Evaluations

Candidate approaches

- ROV – Reflection
- ROV – spatial
- Context-based approach
- YOLO framework (baseline)
- ROV-Empty
- ROV - Tuning

Experimental Evaluations

Dataset

- GeoSpatial object chosen for evaluation is buildings.
- Building datasets are taken from publicly available Massachusetts Buildings Dataset.
- Aerial images of county mosaics from National Agricultural Imagery Program.

Experimental Evaluations

Building dataset

- 127,282 building footprints in the dataset for training the MOBR-based YOLO framework.

- Further 51,326 for testing are used for testing.

NAIP Image dataset

- NAIP image was split into 1022 test images of size 208 x 208.

- Further the test images are categorized into different sub-areas (A1 – A3).

Experimental Evaluations

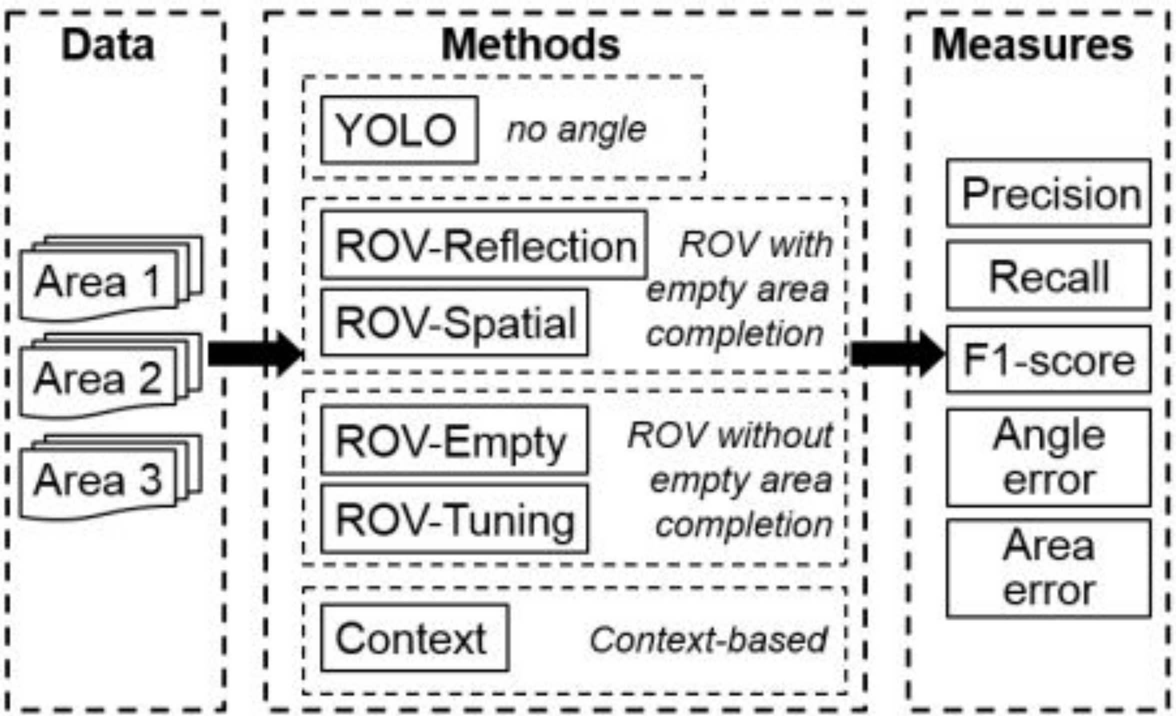
Sub Area	No. Of Images	No. of Buildings
A1	340	17376
A2	341	18954
A3	341	14996

Experimental Evaluations

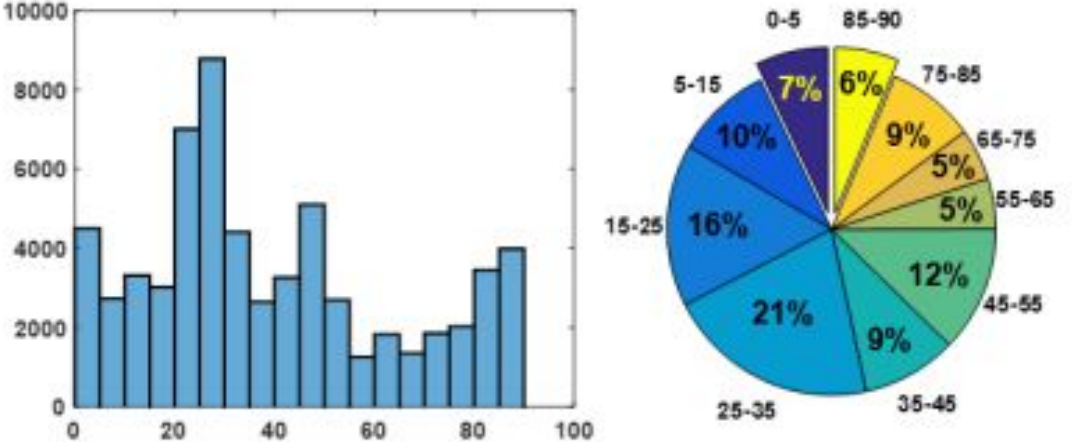
Questions that need to be answered by the experiment

- Do empty areas in rotated images affect solution quality?
- Does ROV-Tuning reduce the effect of empty areas?
- Do the proposed ROV and context-based approaches improve accuracy on angle estimation?
- Do the proposed ROV and context-based approaches improve accuracy on area estimation?
- Do the completion algorithms (ROV-reflection, ROV-Spatial) reduce the effect of empty areas?

Experimental Evaluations



Experimental Results



(a) Distribution of angles (b) Proportion comparison

- Only about 13% of buildings are in orthogonal angle
- For the rest of building when pure MOBRs are used, it leads to overestimation of object areas

Experimental Results

Sensitivity Analysis

- Assess the effect of empty spaces in rotated images
- Results are quantified using Precision,Recall & F1 Score
- The results are compared with YOLO which serves as the basline to measure the effect, since it doesn't use any rotation.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning
A1-Precision	79.8%	81.0%	81.6%	86.4%	83.4%
A1-Recall	69.1%	64.7%	68.1%	52.8%	38.9%
A1-F1 score	74.1%	72.0%	74.2%	65.5%	53.0%
A2-Precision	80.5%	82.4%	83.0%	87.1%	85.3%
A2-Recall	72.8%	68.0%	71.7%	54.9%	44.8%
A2-F1 score	76.5%	74.5%	77.0%	67.4%	58.7%
A3-Precision	78.3%	82.1%	83.2%	87.3%	87.3%
A3-Recall	74.1%	66.8%	71.1%	55.8%	42.5%
A3-F1 score	76.1%	73.6%	76.6%	68.0%	57.1%

Experimental Results

Sensitivity Analysis – Context-based approach

- Three metrics Precision, recall & F1 score is almost same as YOLO.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning
A1-Precision	79.8%	81.0%	81.6%	86.4%	83.4%
A1-Recall	69.1%	64.7%	68.1%	52.8%	38.9%
A1-F1 score	74.1%	72.0%	74.2%	65.5%	53.0%
A2-Precision	80.5%	82.4%	83.0%	87.1%	85.3%
A2-Recall	72.8%	68.0%	71.7%	54.9%	44.8%
A2-F1 score	76.5%	74.5%	77.0%	67.4%	58.7%
A3-Precision	78.3%	82.1%	83.2%	87.3%	87.3%
A3-Recall	74.1%	66.8%	71.1%	55.8%	42.5%
A3-F1 score	76.1%	73.6%	76.6%	68.0%	57.1%

Experimental Results

Sensitivity Analysis – ROV Empty

- Recall is 50% to 55% for areas A1 to A3.
- This is 10% - 20% lower than the baseline YOLO.

- Empty/zero valued pixels could lead to low activation values in deep network layers, and reduce the probability score on objects
- The effect is significant around the border areas between image and the empty areas.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning
A1-Precision	79.8%	81.0%	81.6%	86.4%	83.4%
A1-Recall	69.1%	64.7%	68.1%	52.8%	38.9%
A1-F1 score	74.1%	72.0%	74.2%	65.5%	53.0%
A2-Precision	80.5%	82.4%	83.0%	87.1%	85.3%
A2-Recall	72.8%	68.0%	71.7%	54.9%	44.8%
A2-F1 score	76.5%	74.5%	77.0%	67.4%	58.7%
A3-Precision	78.3%	82.1%	83.2%	87.3%	87.3%
A3-Recall	74.1%	66.8%	71.1%	55.8%	42.5%
A3-F1 score	76.1%	73.6%	76.6%	68.0%	57.1%

Experimental Results

Sensitivity Analysis – ROV Tuning

- lower precision,recall & f1 scores than ROV-Empty
- it is difficult to offset the effects of large chunks of zero pixels.
- kernels which were learned to reduce such effects along the border may have impacted at places with no empty areas around.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning
A1-Precision	79.8%	81.0%	81.6%	86.4%	83.4%
A1-Recall	69.1%	64.7%	68.1%	52.8%	38.9%
A1-F1 score	74.1%	72.0%	74.2%	65.5%	53.0%
A2-Precision	80.5%	82.4%	83.0%	87.1%	85.3%
A2-Recall	72.8%	68.0%	71.7%	54.9%	44.8%
A2-F1 score	76.5%	74.5%	77.0%	67.4%	58.7%
A3-Precision	78.3%	82.1%	83.2%	87.3%	87.3%
A3-Recall	74.1%	66.8%	71.1%	55.8%	42.5%
A3-F1 score	76.1%	73.6%	76.6%	68.0%	57.1%

Experimental Results

Sensitivity Analysis – ROV Reflection & ROV Spatial

- On an average 10% to 15% increases in both recall and F1-Scores compared to ROV Empty
- Proposed algos are able to mitigate the effects of empty areas.
- ROV-Spatial is consistenltly better than ROV-reflection.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning
A1-Precision	79.8%	81.0%	81.6%	86.4%	83.4%
A1-Recall	69.1%	64.7%	68.1%	52.8%	38.9%
A1-F1 score	74.1%	72.0%	74.2%	65.5%	53.0%
A2-Precision	80.5%	82.4%	83.0%	87.1%	85.3%
A2-Recall	72.8%	68.0%	71.7%	54.9%	44.8%
A2-F1 score	76.5%	74.5%	77.0%	67.4%	58.7%
A3-Precision	78.3%	82.1%	83.2%	87.3%	87.3%
A3-Recall	74.1%	66.8%	71.1%	55.8%	42.5%
A3-F1 score	76.1%	73.6%	76.6%	68.0%	57.1%

Experimental Results

Comparative Analysis

- Evaluate the improvements on angle and area estimation by comparing to baseline YOLO
- Results are comapared using Ea = Error of area (%) and Er = Error of rotation angle (°)

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning	Con-text
A1-E_r	26.5°	9.9°	9.1°	11.0°	7.3°	3.5°
A1-E_a	57.2%	29.8%	29.0%	33.2%	32.0%	57.2%
A2-E_r	22.8°	9.2°	8.6°	10.7°	7.3°	3.6°
A2-E_a	48.2%	26.3%	25.5%	28.9%	28.2%	47.8%
A3-E_r	19.0°	10.2°	9.8°	10.9°	9.3°	4.7°
A3-E_a	44.6%	27.2%	26.8%	29.2%	30.0%	44.4%

*Notation: E_a = Error of area (%), E_r = Error of rotation angle (°)

Experimental Results

Comparative Analysis – YOLO (Angle Estimation)

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning	Con-text
A1-E_r	26.5°	9.9°	9.1°	11.0°	7.3°	3.5°
A1-E_a	57.2%	29.8%	29.0%	33.2%	32.0%	57.2%
A2-E_r	22.8°	9.2°	8.6°	10.7°	7.3°	3.6°
A2-E_a	48.2%	26.3%	25.5%	28.9%	28.2%	47.8%
A3-E_r	19.0°	10.2°	9.8°	10.9°	9.3°	4.7°
A3-E_a	44.6%	27.2%	26.8%	29.2%	30.0%	44.4%

*Notation: E_a = Error of area (%), E_r = Error of rotation angle (°)

- YOLO framework models objects with MOBRs, hence it cannot estimate the angles of objects well when thery are not aligned with orthogonal directions of test image.
- Hence Yolo has high errors of angle (about 20° - 25°).

Experimental Results

Comparative Analysis – ROV methods (Angle Estimation)

- ROV methods perform better when compared to YOLO.
- Uses rotation vector with 8 angles, so distance between nearest angles is 11.25°(90°/8).
- If the actual object has an angle right at the middle of two nearby angles, then it will have an error of atleast 5.625° (11.25°/2).
- Hence, it may partially contribute to the 5° error gap.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning	Con-text
A1-E_r	26.5°	9.9°	9.1°	11.0°	7.3°	3.5°
A1-E_a	57.2%	29.8%	29.0%	33.2%	32.0%	57.2%
A2-E_r	22.8°	9.2°	8.6°	10.7°	7.3°	3.6°
A2-E_a	48.2%	26.3%	25.5%	28.9%	28.2%	47.8%
A3-E_r	19.0°	10.2°	9.8°	10.9°	9.3°	4.7°
A3-E_a	44.6%	27.2%	26.8%	29.2%	30.0%	44.4%

*Notation: E_a = Error of area (%), E_r = Error of rotation angle (°)

Experimental Results

Comparative Analysis – Context based (Angle Estimation)

- As Expected contex-based approach performs best.
- Recall, that context information is available, to calculate the angles.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning	Con-text
A1-E_r	26.5°	9.9°	9.1°	11.0°	7.3°	3.5°
A1-E_a	57.2%	29.8%	29.0%	33.2%	32.0%	57.2%
A2-E_r	22.8°	9.2°	8.6°	10.7°	7.3°	3.6°
A2-E_a	48.2%	26.3%	25.5%	28.9%	28.2%	47.8%
A3-E_r	19.0°	10.2°	9.8°	10.9°	9.3°	4.7°
A3-E_a	44.6%	27.2%	26.8%	29.2%	30.0%	44.4%

*Notation: E_a = Error of area (%), E_r = Error of rotation angle (°)

Experimental Results

Comparative Analysis – (Area Estimation)

- ROV methods were able to reduce the area erros from avg 45%-55% to 25%-30%.
- Context based approach only focuses on angles, hence it is not able to reduce the errors on area.
- ROV-Spatial shows better performance when compared to others.

Area ID-Metric	YOLO	ROV-reflect.	ROV-spatial	ROV-empty	ROV-tuning	Con-text
A1-E_r	26.5°	9.9°	9.1°	11.0°	7.3°	3.5°
A1-E_a	57.2%	29.8%	29.0%	33.2%	32.0%	57.2%
A2-E_r	22.8°	9.2°	8.6°	10.7°	7.3°	3.6°
A2-E_a	48.2%	26.3%	25.5%	28.9%	28.2%	47.8%
A3-E_r	19.0°	10.2°	9.8°	10.9°	9.3°	4.7°
A3-E_a	44.6%	27.2%	26.8%	29.2%	30.0%	44.4%

*Notation: E_a = Error of area (%), E_r = Error of rotation angle (°)

Visualization

(a) YOLO (b) ROV-Reflection (c) ROV-Spatial (d) ROV-Empty (e) ROV-Tuning (e) Context

Future Work

- One limitation of current ROV scheme is that it cannot cover all possible angles due to the discrete angle in rotation vector.
- Future work, is to explore new technique to further refine angles without adding much computational overhead.
- Explore an integrated approach fo combinging ROV and context-based method to improve the solution quality.