Al-Azhar Engineering 16th International Conference

Vol. 19, No. 72, July 2024, 156 - 172

APPLICATION OF UAV DATA AND GEOSPATIAL AI TECHNIQUES FOR SEWER INLETS LOCALIZATION AND MAPPING.

Haysam M. Ibrahim*, Essam M. Fawaz, Amr M. Elsheshtawy, Ahmed M. Hamdy

Department of civil engineering, college of engineering, Al-Azhar University, Cairo, Egypt

***Correspondence:** Hythammahmoud.14@azhar.edu.eg

Citation:

H.M. Ibrahim, E.M. Fawaz , A.M. Elsheshtawy and A.M. Hamdy," Application of UAV data and geospatial ai techniques for sewer inlets localization and mapping", Journal of Al-Azhar University Engineering Sector, vol. 19, pp. 156 - 172, 2024.

Received: 02 December 2023

Revised: 17 January 2024

Accepted: 24 January 2024

Doi: 10.21608/auej.2024.255068.1519

ABSTRAT

Unmanned aerial vehicle (UAV) systems underwent significant advancements in recent years, which enabled the capture of high-resolution images and accurate measurements, with the tremendous development in artificial intelligence, especially deep learning techniques, Which allows it to be used in the development of Drainage infrastructures that represent a major challenge to confront the flood risks in urban areas and represent a considerable investment, but they are often not as well classified as they should be. In this study, we present an automatic framework for detecting, localizing, and mapping sewer inlets from image clouds acquired by UAVs based on a YOLO CNN architecture. The framework depends on the high image overlap of unmanned aerial vehicle imaging surveys, which were then processed using Structure-from-Motion (SfM) to generate orthomosaic imagery. The framework uses a YOLOv5 model trained to detect and localize sewer inlets in aerial images with a ground sampling distance (GSD) of 3 cm/pixel. Novel object-detection algorithms, including YOLOv5n, YOLOv5s, and YOLOv5x, were compared in terms of the classification and localization of sewer inlets. The approach is evaluated by cross-validating results from an image cloud of 250 UAV images captured over a 0.57 km^2 study area with 228 sewer inlets. Images with models' performances from the literature, the new YOLO model tested on UAV images in this study demonstrates satisfactory performance, improving both precision and recall. The results show that YOLOv5x offers the best precision (90%) and recall (92%), whereas YOLOv5n achieved less accuracy in precision and recall (78%) and (80%), respectively. Additionally, increasing image size in the training stage is a very important modification in the model. The study approach has a remarkable ability to detect sewer inlets and can be used to develop the inventory of drainage infrastructure in urban areas.

KEYWORDS: Drainage mapping, YOLO algorism, small object detection, unmanned aerial vehicle, urban drainage.

تطبيق الطائرات بدون طيار وتقنيات الذكاء الاصطناعى لتوطين مداخل صرف المطر ورسم الخرائط

هيثم محمود إبراهيم محمود*،عصام محمد عبد الباسط فواز،عمرو محمود الششتاوى، أحمد محمد حمدى

قسم الهندسة المدنية، كلية الهندسة، جامعة الازهر، مدينة نصر، 11884، القاهرة، مصر.

*** البريد الإلكتروني للباحث الرئيسي:** Hythammahmoud.14@azhar.edu.eg

APPLICATION OF UAV AND GEOSPATIAL **AI** TECHNIQUES
FOR SEWER INLETS LOCALIZATION AND MAPPING.

الملخص

شهدت أنظمة الطائرات بدون طيار تطورا هائلا في الآونة الأخيرة لتوفر صوروقياسات عالية الدقة, بالإضافة للتطور الهائل في أنظمة الذكاء الاصطناعي خاصة تقنيات التعلم العميق. يمكن استخدام هاتين التقنيتين في تطوير البنية التحتية لصرف المطر التي تمثل تحديا كبيرا في مواجهة مخاطر الفيضانات في المناطق الحضرية وتمثل استثمارا ضخما لكنها في كثير من الأحيان لا يتم تصنيفها كما ينبغي. في هذه الدراسة نقدم اطارا اليا للكشف عن مداخل صرف المطر وتحديد موقعها ورسم خرائط لها من الصور التي تم الحصول عليا بطائرة بدوت طيار واعتمادا على بنية الشبكات العصبية في نموذج (YOLO) , يعتمد الإطار على التداخل العالي للصور الجوية والتي تمت معالجتها باستخدام خوارزمية (SFM) لإنشاء الصورة (Orthoimage). يستخدم الإطار نموذج (YOLOv5) الذي تم تدريبه لاكتشاف وتحديد مواقع مداخل صرف المطر في الصور الجوية المأخوذة بدقة 3 سم/بكسل. تمت مقارنة خوارزميات الكشف الجديدة لكل من (YOLOv5n, YOLOv5s, and YOLOv5x) من حيث تصنيف وتوطين مداخل صرف المطر. يتم تقييم هذا النهج من خلال التحقق من صحة النتائج من سحابة الصور المكونة من 250 صورة تم التقاطها بطائرات بدون طيار على مساحة أكبر من نصف كيلو متر مربع يوجد بها 222مدخل صرف مطر. بالنظر الى الصور مع أداء النماذج في الأبحاث السابقة يوضح النموذج الجديد لهذه الدراسة الذي تم اختباره على صور الطائرات بدون طيار أداء مرضيا حيث يحسن الدقة والاستدعاء وتظهر النتائج ان (YOLOv5x) يوفر دقة 90% واستدعاء 92% وله قدرة عالية في اكتشاف مداخل صرف المطرورسم خرائط لها .

الكلمات المفتاحية : رسم خرائط صرف المطر، خوارزمية YOLO، كشف الأجسام الصغيرة، طائرات بدون طيار، صرف المطر للمناطق الحضرية.

1 Introduction

Urban growth is a continuing trend, and the development of entombed utility networks is an important part of its outgrowth, but locating network items is a hard mission. The object classification (larger than houses or equal) in aerial images has been extensively researched in the last studies, but small object localization is very challenging and rarely studied because of many factors, such as the difference in object colors, crowded neighborhoods, aspect ratios, shadow effect, and non-uniform background [1].

Artificial Intelligence (AI) and Remote sensing (RS) present an automated solution to error-prone and expensive traditional data collection. (RS) processes depend on data collected for small infrastructure objects at street level that are found on roads. In that respect, aerial imaging is very reliable; aerial imagery with high resolution can normally be georeferenced and rectified with centimeter-level accuracy. Although the georeferencing accuracy and rectification depend on the quantity and quality of GCPs, this is usually not important in height terrain urban areas. UAV imagery achieves this, particularly thanks to its large image overlap and very high GSD.

Several studies investigated manhole cover detection, but sewer inlet (S.I) detection in aerial imagery has not been studied. From a standpoint of remote sensing, sewer inlets are like manhole covers in terms of construction material, frequency, location of occurrence, and size. So, it's important to illustrate the latest research in the detection of manhole cover. (Pasquet 2017) combined detection method, predictions from the SVM support vector machine, and HOG features with detection from a geometric circle filter. Trained and tested on GSD 4 cm/pixel aerial imagery, the method achieved detection equal to 40% of total manholes with 80% precision [2]. More recently, (Commandre 2017) implemented a deep learning CNN to detect manholes using aerial imagery with a (5 cm/pix) resolution. although the resolution is lower, the performance achieved a recall of 50% and a precision of 69% [3]. The most of studies apply detection in single view, to enhance the performance of detection, (Vitry, Schindler et al. 2018) using multiple views, and a (Viola and Jones) model trained to detect (S.I) in aerial images with 3–3.5 cm/pixel resolution. In the Multiview approach compared with the single-view detector, results showed improvement in average precision, which increased from 65% to 73% [1]. The proposed method by (Zhou 2022) Beginning, coarsely classifies the images into

types of rainy and non-rainy then based on the coarse classification results performs manhole cover detections. the method achieves an accuracy to detect manhole cover of 86% and F1-score of 87% using the SVM model [4].

For aerial detection applications, (UAVs) are a natural competitor: when creating orthoimage using UAVs, recommended high overlap captured aerial images to reduce perspective distortion resulting from low flight height. It is recommended, to have an overlap of more than 50% in the side and 70% in the front as shown [5] The high accuracy achieved at (50 m) a flight altitude and (10) GCPs. also, proves that the accuracy of (X, Y) depends on the GCPs number. (Mirko 2019) illustrated the effect of the GCPs on the geometric accuracy of photogrammetric. The results illustrated that 3 GCPs for georeferencing for GIS applications, recommended that 7 GCPs and cartographic production need 15 GCPs [6]. (Jiménez-Jiménez 2021) showed the DTMs' accuracy and quality depend on four factors: (1) the UAV system, either camera or UAV platforms; (2) image acquisition and flight planning (image overlap, flight altitude, flight line orientation, UAV speed, georeferencing, and camera configuration) (3) photogrammetric digital terrain model generation (software, ground filtering, and DEM generation) (4) geomorphology.

They recommended optimizing process variables to have high-accuracy DTMs [7]. (Liu, Han et al. 2022) studied the influence of five factors on UAV photogrammetry (altitude, image resolution, overlap (side and front), GCPs, and focal length), The results prove a better design for processes that produce high accuracy [8].

Since a few years, (J.Redmon 2016) presented a novel model based on a Convolution Neural Network for object detection named You Only Look Once (YOLO). They use a single CNN to predict bounding boxes and then class probabilities directly from the full image in one evaluation. So, the detection pipeline can be enhanced directly completely on prediction performance [9]. Since the original version (YOLOv1) of the novel model achieves a low (mAP), they have improved the model to a new version (YOLOv2) [10].

YOLO deep learning method in the road field presents an applicable method to the detection of road diseases like pavement crack prediction as illustrated [11]. (Zhu 2022) proposed the capture of road distress images using a UAV. depend on three CNN models for object detection YOLOv4, R-CNN, and YOLOv3, were trained on the dataset, and compared their performance [12]. Agriculture field, (Puliti and Astrup 2022) used the YOLOv5 model to recognize tree leaves and then classify trees depending on the leaves damage. UAV imagery was acquired from 89 study areas and was manually annotated into 55 thousand single trees classified into three classes based on their health. The results showed a precision of 76 % and recall of 78 % [13].

Recently, new versions of the YOLO model, namely, YOLOv5n, YOLOv5s, and YOLOv5x, have become more accurate in object detection [14].

1.1 Scope and Novelty of Study

The study aims to create an automated method for detecting small objects to identify sewer inlets using UAV clouds for orthophoto. In total, 250 images were acquired covering an area =0.57 square kilometers used as a case study with 220 S.I sewer inlets, Additionally, the optimal UAV flight settings were examined to ensure accurate orthophoto and DEM generation using the Agisoft Metashape program. Novel Object-detection algorithms, including YOLOv5n, YOLOv5s, and YOLOv5x were compared the classification of sewer inlets. This study is the first clarification of the latest YOLO detector based on UAV images in the water management field.

2 Datasets and Study Area

2.1 Data Acquisition

A low-cost UAV (DJI Phantom) is used to acquire images during a half-hour maximum flight time, The UAV has a 16 MP compact digital camera (8.8 mm Focal Length) controlled by UAV autopilot. The GCPs coordinates were determined by a Trimble R8 GNSS. The UAV was flown over an area of 0.57 km^2 and 90 m flight height. In total, 250 images were taken with a GSD of 3.5 cm/pixel UAV data source (vitry, et al 2018), and our results regarding orthoimage, Dem generation, and their errors against ground control points are considered identical. as shown in **Table. 1**.

Table. 1 The study area Characteristics, and UAV.

Location	Zurich, Switzerland
weather case during flight	Overcast
area	0.5 square kilometer
Altitude	Average =90 m
Lateral and Frontal overlap percentage	60% and 70% respectively
data quantity	250 images
Acquisition data duration	2*0.5 hr.
GSD Image	3-3.5 centimeter per pixel
Resolution	4864 x 3648 pixels
No. of GCPs	10

2.2 Image processing and orthophoto production

The Agisoft program was used for image processing [15] to evaluate the external and internal parameters of the camera and produce an orthoimage **Fig.1** as well as DEM for the case study using the steps as shown in the program. The orthoimage is generated from the image projections by making a mosaic, for orthoimage georeferencing ten GCPs were used. The study area is shown in **Fig.1** and the registration error is shown in **Table. 2** The processing time to generate the orthophoto and DEM was 12 hours (using core i5, 12 GB RAM, and card 2 GB).

Table 2. Registration error of DEM and Orthoimage Generation evaluated at (GCPs).

GCP	Error X (cm)	Error Y(cm)	Error Z(cm)
Min	-1.9	-2.5	-6.4
Max	2.3	3.3	5.9
Mean (cm)	0.03	0.13	-0.59
Sigma (cm)	1.32	1.99	2.98
Root Mean Squar Error (cm)	1.32	2.00	3.04

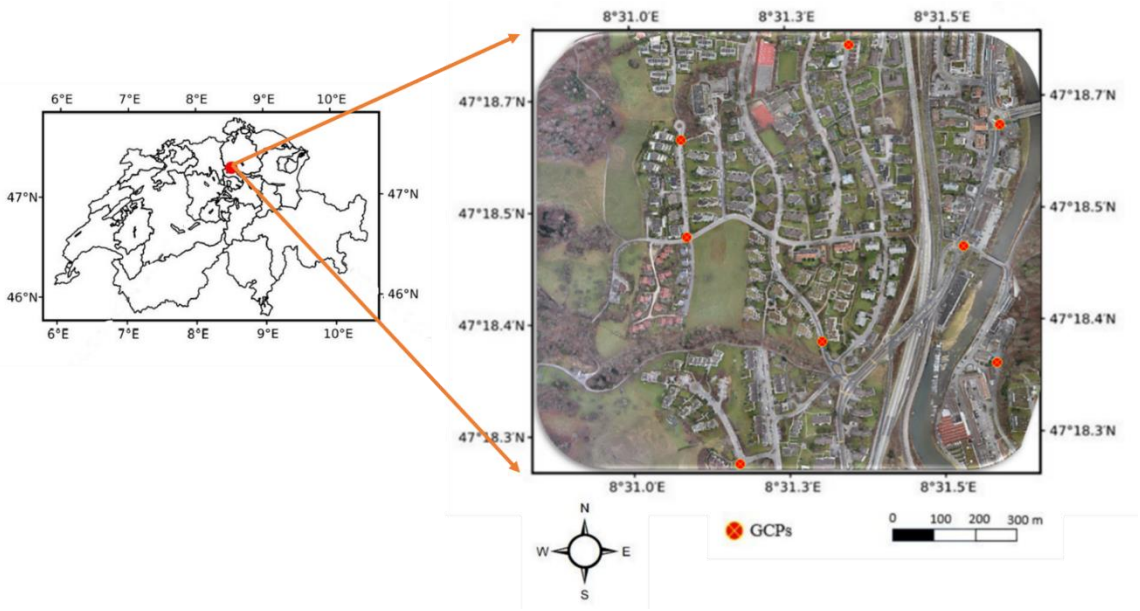


Fig.1. The Case Study (0.57 km^2) Zurich, Switzerland. The Generated Orthoimage and DEM.

3 Experiment

Table . Data collection instruments and image processing tools.

Table 3. Methodology list:

Seq	Tool	Function
1	UAV DJI-phantom	Image acquisition.
2	GPS	Observation GCPs coordinates.
3	Metashape Agisoft	Orthophoto and DEM Generation.
4	Python	The model writing language.
5	Roboflow	Dataset annotation.
6	YOLO Detector	Automatic Sewer inlets Localization.
7	Google Colab	Environment of Model Training and Validation.
8	CPU	Environment of Model testing.
9	Arc GIS	Map Production & Orthophoto Clipping and Merging.
10	ANOVA	Analysis of Variances for results.

3.1 Experimental environment

The YOLOv5 detectors were trained using 250 full-scale images captured with a resolution of 4864 x 3648 pixels, which needs a long processing time to train detectors. Because of the limited computing resources, the experiments were conducted on a Google Colab cloud server, as shown in **Table 4**.PyTorch 1.2.0 was the experimental frame.

Python was used for writing object-detection algorithms.

To accelerate training GPUs were used.

Table 4 details of google colab server:

NVIDIA- 525.85.12	Driver: 525.85.12	CUDA:12.0
GPU Name Fan Temp Perf	Bus-Id .A Memory-Usage	Volatile .GPU- M.
Tesla T4 N/A 45C P0 25W/70W	0MiB / 15360 MiB	0% Default N/A

3.2 Framework of Machine learning

Workflow to train YOLO, as illustrated in **Fig.3,** The study depends on resizing original image into two image sizes before input to network. First, the original image was resized to 640 pixels in YOLO5n, YOLO5s, and YOLO5 x.

Second, the original image was resized to 1280 pixels in YOLO5n6, YOLO5s6, and YOLO5 x6. Before the training stage, Each Sewer Inlet was labeled manually by rectangular boxes (ground-

truth boxes) in all images. The collected images were divided into a training dataset with proportions of 0.8 and a validation dataset with proportions of 0.2.

Data augmentation using roboflow was performed to increase the training dataset by three times. Finally, the dataset for training based on epochs (100 to 300) was input to the YOLO network architecture.

3.3 Model training

The YOLOv5 Convolution Neural Network architecture is a single-stage detector consisting of three parts:

The first component is named backbone to extract features of an input image using cross-stage partial architecture, that enhanced model size and speed of boosting [16], [17].

The second component named the neck is used to enhance ability the model on object scaling and unseen data transformation using path Aggregation net [18].

The third and final component named Head, output vectors generated here with the confidence of class, bounding box, and scores. such a step is important to perform the final detection parameters. This step-in version 5 and previous versions have not changed [13].

The Sewer Inlets prediction using the YOLOv5 model was studied in terms of precision, Recall, and F1 score. **Fig. 2** shows the small objects (Sewer inlets) used in training.

Fig. 2 The type of small objects (S.I) used in Experiment.

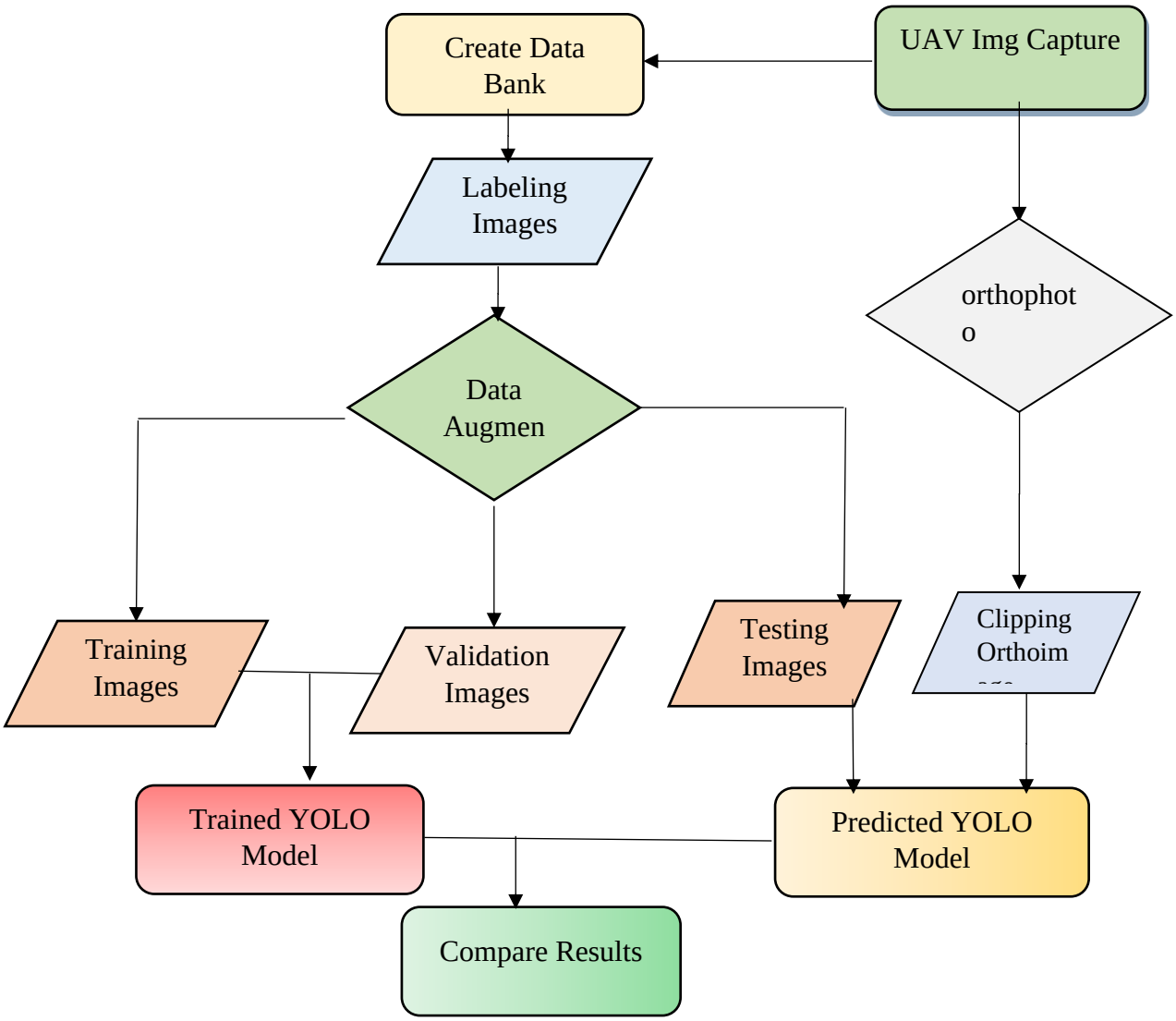


Fig.3. flowchart of the proposed methodology.

3.4 . Loss function

YOLOv5 model corrected confidence values and coordinates of the bounding box using the square error loss it is called the loss function [9]. Regression of the bounding box, objectness, and probability of a class, for the three scores, computed the loss function then the compounded loss function [13].

3.5 Evaluation metrics

Evaluation metrics of the model are Recall (R), Precision (P), which is equivalent to accuracy, and intersection over union (IoU) identical to mean average precision (mAP) [13] as illustrated in Eq. (1) and Eq. (2):

$$Accur = \frac{TN+TP}{TN+TP+FN+FP} \qquad \text{Eq. (1)}$$

$$IOU = \frac{(Bp \cap Bb)}{(Bp \cup Bb)} \qquad \text{Eq. (2)}$$

Where: B_p = Bounding Box predicted, B_g = ground truth (BB),

(TP) = number of correctly predicted object samples,

(FP) = predictions number where non-objects are founded,

(FN) = number of undetected objects,

Accuracy model evaluated by metrics, P, R, and F1 as shown in **Table** and Equation. (3), (4) and (5) [19]:

Table 5 The bounding box identifications:

Ground Truth bounding box.	Results Prediction	
	(Positive)	(negative)
Objects (existing actually)	(TP)	(FN)
Nonobjects (not existing actually)	(FP)	(TN)

$$Recall = \frac{TP}{TP+FN} \quad \text{Eq. (3)}$$

$$Precision = \frac{TP}{TP+FP} \quad \text{Eq. (4)}$$

$$F1 = 2\frac{Precision.Recall}{Precision+Recall} \quad \text{Eq. (5)}$$

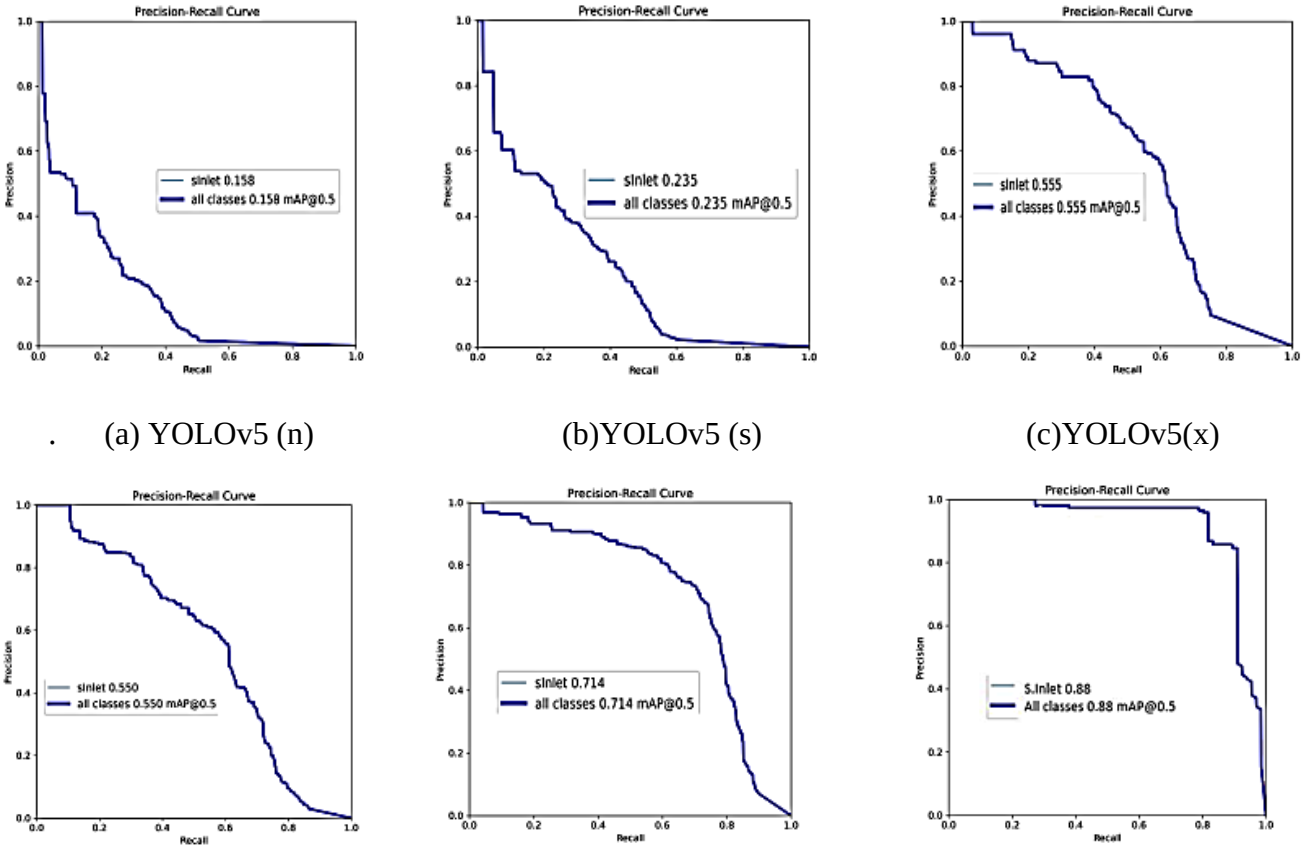
Precision score equals correct prediction percentage, A greater IOU indicates a smaller P. Recall score equals all positives that is the model predicted. The F1 score indicates confirmation of R and P metrics, and it is the harmonic mean of them. AP equals the P–R curve integration, which equals the area under the P_R curve.

3.6 post-processing and mapping

The possibility of using YOLOv5n, YOLOv5s, and YOLOv5x for purposes of mapping with orthophoto generating by image UAV, needs to add post-processing steps, as follows: First The UAV orthomosaic Tiling: As described by [20] the orthomosaic were tiled into suitable area geotiff tiles (The image size in the prediction stage equals the image size in the training stage) using an overlap = 2 m between tiles. Second Predicting each tile: using the best weights generated from the training model, for each tile was predicted coordinates of bounding box. Third, Image/map: YOLOv5 predicted bounding box coordinates from the origin images top left corner in pixel values and then we converted it using the UTM coordinates to a geographical space. Fourth, Bounding/box: when the intersection between bounding boxes pairs was > 0.75, the smallest confidence of class and box is discarded.

4 Results and Discussion

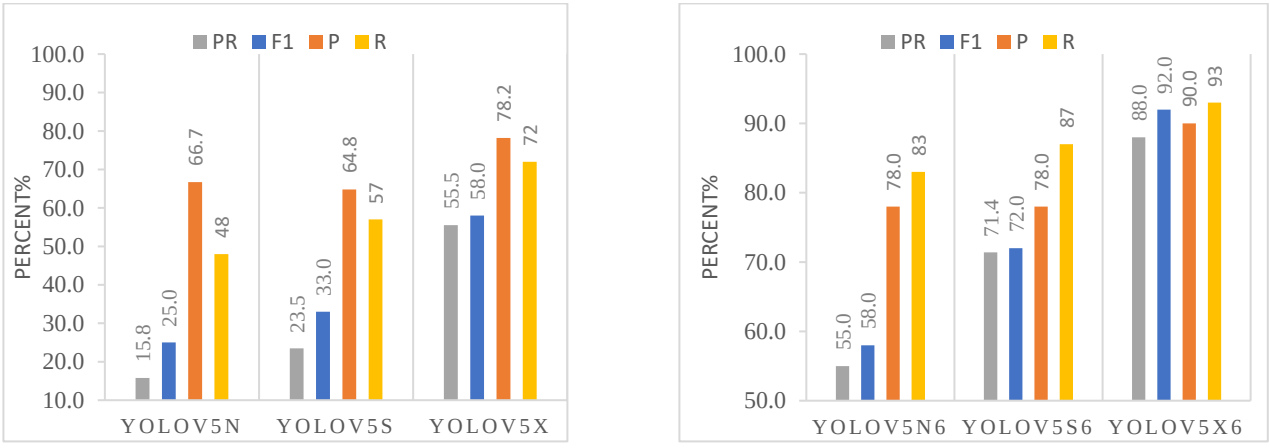
4.1 YOLOv5 detectors and Sewer Inlets



. (a) YOLOv5 (n) (b)YOLOv5 (s) (c)YOLOv5(x)

(d) YOLOv5 (n6) (e)YOLOv5 (s6) (f)YOLOv5 (x6)

Fig. 4 Sewer inlet Precision Recall curve.



(a). Img Size= 640 (b) Img Size -1280.

Fig. 5. Metrics (P-R, P, R and F1 score) for three YOLOv5 detectors.

Fig. 4. illustrates the P–R curves to predict the sewer inlet by three different YOLOv5 architectures. **Fig. 5** (a) shows the Sewer inlets P, R, PR, and F1 score using the YOLOv5 detectors.

YOLOv5n exhibits inferior performance in sewer inlet detection, with a mAP =15.8%. YOLOv5s exhibit a better performance than YOLOv5n with F1 score and recall of 33% and 57%, respectively. YOLOv5x demonstrated better overall performance, for a metric of an F1 score equal to 58% and a Recall equal to 72%.

Fig. 4 and **Fig. 5** (b) show the Sewer inlets P, R, PR, and F1 score using the YOLOv5 Modified detectors (n6, s6, x6), results show better performance when the image size was modified in the training model from 640 to 1280.YOLOv5n performance in sewer inlet detection, a mAP increased from 15.8% to 55%. YOLOv5s6 achieved a better performance than YOLOv5s with F1 score and recall from 33% and 57% to 72% and 87% respectively. YOLOv5x6 demonstrated better overall performance, with an F1 score from 58% to 92% and a Recall from 72% to 93%.

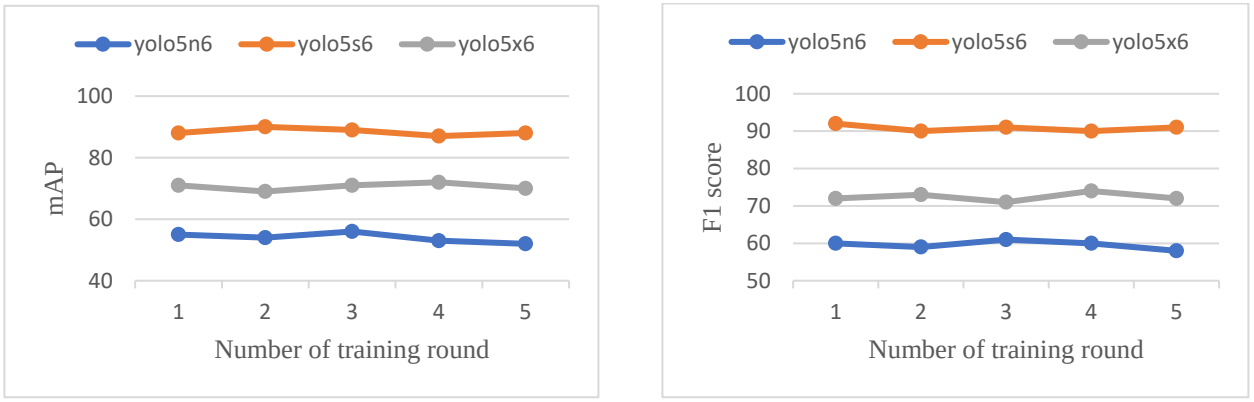
Vitry.2018 used the same 252 UAV images for S.inlet detection, achieving (AP) = 73% with a Viola–Jones classifier [1]. [21] used 400 UAV images to a training data set for pavement distress detection, the results were (mAP) = 53.1% with the YOLOv3 architecture. [22] conducted changes on YOLOv3 architecture to improve the performance of the original baseline, where the results for (mAP) were 52.3%. [23] used the YOLOv3 tested on UAPD for pavement distress detection, achieving (mAP) = 56.6%.

Fig. 5 shows that YOLOv5x6 and YOLOv5s6 generally performed better than yolov5n6. Anchors that are used can extract the features of an object in all regions, through higher features of semantics and the sequence of low-level information of texture, YOLOv5n, yolov5s, and YOLOv5x are strong enough to detect sewer inlets.

YOLOv5x6 has better performance and a higher mAP, vs. poor performance in YOLOv5n to predict Sewer inlets. More than that, with studied deeply can achieve better mAP value.
The experiment was repeated on the three model versions to prove the model stability in the Sewer inlet prediction using this dataset.

4.2 Effect of repeat training

Five additional training experiments were performed to ensure the validity of the model prediction performance. The experimental configuration was the same as that in the previous training. **Fig. 6** (a) and (b) shows the mAP and F1 score of the third models. **Fig. 6** illustrates in the general direction, the five experiments results were stable performance. The error is in the range of 3% between all experiments, with volatility, it is common. Although the same weights are used, the training results are slightly different each time. Due to the random selection of training images in each batch and the variety of gradients, it led to slightly different predictive results in each training round. The results of the YOLOv5x6 model once again showed a performance distinct from other models.



(a) (b)

Fig. 6. Model prediction with repeated training (mAP) and (F1 score) respectevly.

4.3 Effect of MaxEpoch on dataset augmentation

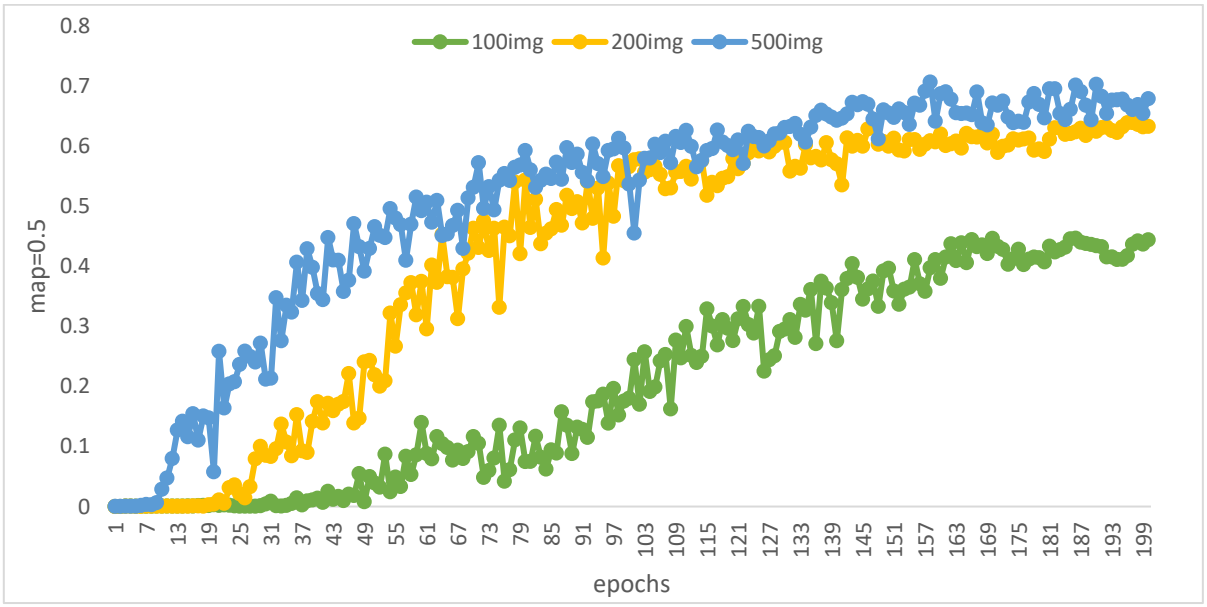
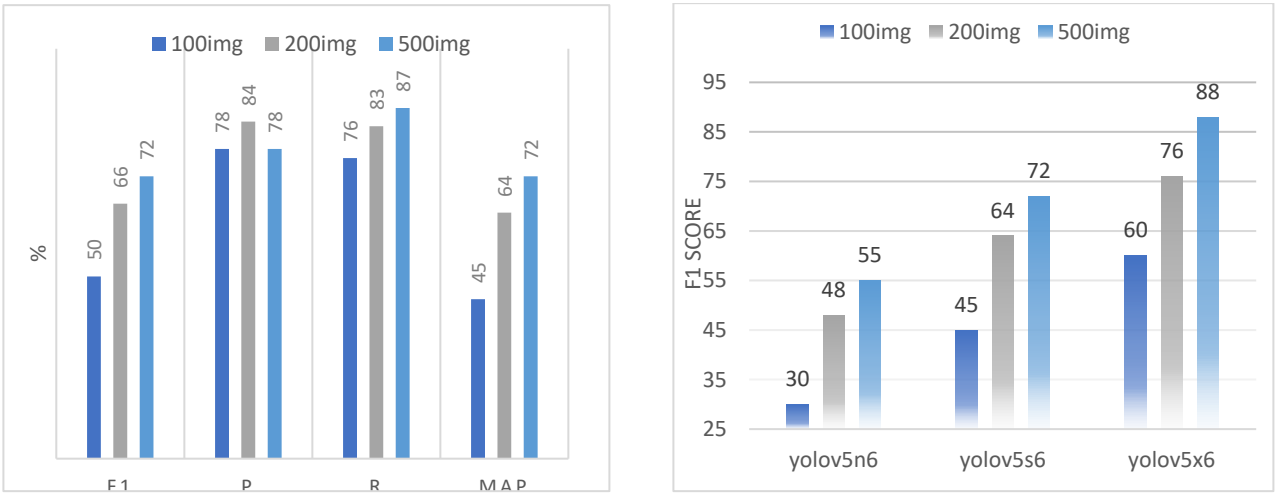


Fig. 7. Effect MaxEpochs on Dataset augmentation based on YOLOv5s6.

The accuracy of Sewer Inlet prediction significantly depends on MaxEpoch adjusting. for suitable performance to the Sewer inlet detection using YOLOv5 detectors, MaxEpoch should be at least 100. An epoch refers to the total number of repeats for applying training on all images used in one round, it is important to deep learning performance [19]. Identifying the MaxEpochs numbers can improve model architecture performance to get a better prediction result, after that, no significant performance enhancement. This study used max iterations of 200 in the training stage. **Fig. 7.** shows the YOLOv5s6 testing metrics results to discuss its effects. When MaxEpochs increased up to 100, the model performance accuracy increased significantly for object detection. Increasing MaxEpoch by more than 100 did not achieve significantly better accuracy, in addition increased training time in case 200 images and 500 images, but it increased after 150 MaxEpochs in case 100 images. Based on the study results, the suitable MaxEpoch ranged between 100 and 200 for achieving accuracies exceeding

72% by 500 images. in general dataset augmentation achieved higher (72%) accuracies at less than 150 MaxEpoch. indicated to augmentation dataset reducing MaxEpochs.



(a) (b)

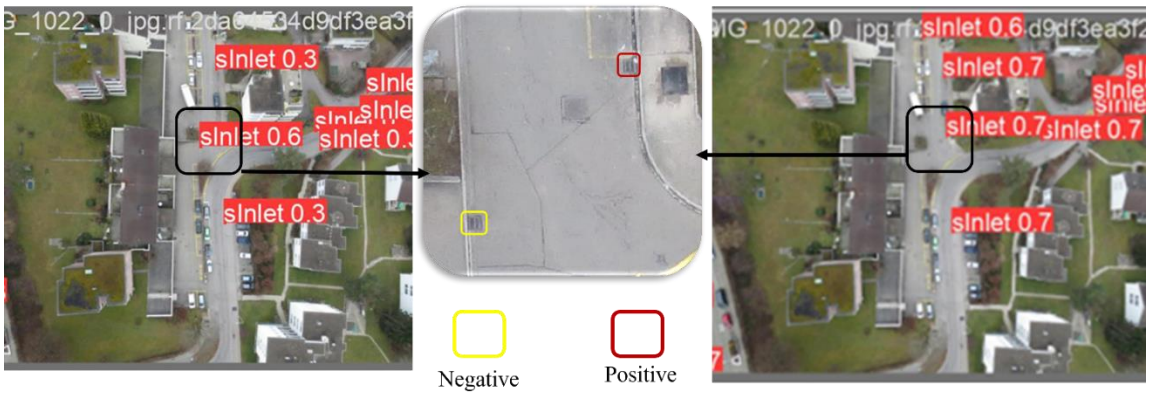
Fig. 8. Effect of Dataset augmentation.

The study presents the F1 score, Precision, Recall, and mAP for testing datasets, as shown in **Fig.** 8 **(a).** yolov5s6 achieved values (72%, 84%, 87%, and 72%) for the F1 score, Precision, Recall, and mAP. The reliability of autonomous Sewer Inlet detection in UAV images using convolution neural networks is dependent significantly on the size of the training and validation dataset. Through the training stage, 100, 200, and 500 training images were prepared with augmentation and then entered the network of YOLOv5. **Fig.8** illustrates the results of YOLOv5 conducted using different datasets. Overall, the prediction metrics of Sewer Inlet increased with increasing the number of images.

YOLOv5s6 achieves a better performance in accuracy, (Recall was increased from 76% to 87%) when augmentation was applied to training images. As shown in **Fig. 8 (b)** YOLOv5n6 showed relatively low accuracy to predict Sewer Inlet, When the dataset number was increased from 100 images to 500 images, mAP increased to 55%. On the other hand, YOLOv5x6 gives the best accuracy. It is recommended, based on this evaluation, when training a YOLOv5 for reliably predicted Sewer inlets in UAV images, must use a minimum of 500 images.

4.4 Results of YOLOv5 models

Results of YOLOv5 model as shown in **Fig. 9.(a)**. Predictions of Sewer inlets using YOLOv5x were generally better than YOLOv5s. Based on the mAP of the (S.I) was calculated as 75 %, for the three models. Sewer inlets detection using YOLOv5x6 was 90% for Recall, compared with 73% achieved by (Vitry, 2018), proof of the potential of UAV sewer inlet prediction based on CNN. Sewer inlet prediction results by YOLOv5s6 and YOLOv5x6 are presented in **Fig. 9.(a)** and **Fig. 9.(b)**, YOLOv5x6 is the best model in prediction. Sewer inlets have parallel features and are extensively annotated in a dataset, so, can be predicted well with high-confidence values as shown in **Fig.9 (a) and (b)**.



(a) YOLOv5s (b) YOLOv5x

Fig. 9. Sewer inlets detection results.

4.5 Orthophoto Predictions

The YOLOv5 model appeared excellent consistency between the visible Sewer Inlets and Bounding box prediction, and highlighted, in the Orthophoto as shown in **Fig. 10.** The model has an important advantage as the ability to recognize Sewer Inlets in crowded scenes and different features in Orthophoto with a (GSD) of 3.5 cm/pixel, the yolov5 model achieved 90% accuracy for Sewer Inlets detection.

Fig. 10. Prediction of sewer inlets in Orthophoto after clipping.

4.6 ANOVA Two-way test

To analyze the model accuracy among Image sizes, ANOVA (two-way) is used [24] ,[25] ,[26] . The test is conducted to compare the (mAP) for two image sizes in each model. Two versions of the YOLO model (V5s and V5x) called M and two image sizes (640 and 1280) called S are taken as factors. ANOVA first step, should be declared the null, and the alternative hypothesis [27]. A Null hypothesis is called H0. The alternative is called H1.

H0: $X_{M1} = X_{M2} = X_{M3}$: $:X_{Mn}$ (means of factor M is equal). Eq.(6)

H0: $X_{S1} = X_{S2} = XS3$: $:X_{Sn}$ (means of factor S is equal). Eq. (7)

H0: $X_{MS1} = X_{MS2} = X_{MS3}$: $:X_{MSn}$ (means M and S are equal) Eq. (8)

ANOVA was performed as illustrated in [28]. Here, the sum of squares was declared as three types: sum of square for factor M (SSM), sum of square for factor S (SSC), and sum of square for both factors M and S (SSMS).The two-way ANOVA can be calculated by declaring the factors M and S as model and class, respectively. The result of the calculation is shown in **Table 6.** From **Table 6**, the F-value is calculated as 130. So, it is denoted as calculated F-value.

F -value (calculated) =70 image size Eq. (9)

F -value (calculated) =1.43 model Eq. (10)

Critical F-value is 4.11. So, it is denoted as tabulated F-value.

F -value (crit)= 4.11 Eq. (11)

Table 6. ANOVA Results:

Source of Variation	*SS*	*df*	*MS*	*F*	*P-value*	*F crit*
Image Size	20702.5	1	20702.5	70.24411	5.5577E-10	4.113165277
Models	422.5	1	422.5	1.433553	0.23901009	4.113165277
Interaction	562.5	1	562.5	1.908577	0.17563605	4.113165277
Within	10610	36	294.7222			
Total	32297.5	39				

Image size factor, the calculated F-value is 70 and the critical F-value is 4.11. Hence, the null hypothesis can be rejected, and the alternative hypothesis can be accepted.

Models' factor, the calculated F-value is 1.43 and the critical F-value is 4.11. Hence, the null hypothesis can be accepted. Analysis proves that there is a significant relationship between the mean values of the S factor, but there is no significant relationship between the mean values of the M factor.

By declaring the models and Image size as two factors, the two-way analysis proves that the mean values of (mAP) vary between each image size, and the mean values of (mAP) no more change between each model. (mAP) was better During the 1280 image size, it reached 88% with yolov5x, but in 640 image size, it was less than 25% with yolov5s.

5 Recommendation

There are many possibilities for the development of an automated sewer inlet detection approach and drainage mapping using aerial imagery, but two main limitations need more study.

First limitations, the sewer inlets are not visible in (UAV) images because they are temporarily covered by trees and debris or vehicles. This problem can be reduced by performing flights, at different times.

The second limitations, there are many sewer inlet forms and situations, in our case study only one type of sewer inlet is trained. To adjust this variety , first must increase the training data variety, second, best adaptation of image capturing, like increasing camera tilt. Therefore, depending on the relevance of the illustrated limitations and the completeness required of the data, it may be necessary either adjust the detection approach or verify the detection results manually.

Supplementary Materials: The data used in this paper is available online at https://zenodo.org/record/1197592 ,

6 Conclusion

The paper studied how to develop a YOLOv5 model, through the combined use of a labeled database of UAV images and Convolution Neural Networks, to predict sewer inlets accurately and mapping it. The important study results are that the model achieved suitable accuracy metrics for a sewer inlet, with a validated case study using 250 UAV full-scale images.

Results show that the use of the YOLOv5 model increases (AP) from 0.73 to 0.90 as compared to a literature review using the same GSD of the dataset. The gain is attributed to the ability to exploit the full resolution of the raw UAV images.

This study can identify more than 91% of the (S.I) with a precision of 90% and localize them when using a YOLOv5x6 detector that trained on 250 UAV images and max epochs 200 and based on YOLOv5n at the same factors was able to identify about 80% of the (S.I) with precision of 78%.

Additionally, this study illustrated that increasing image size from 640 to 1280 in model training, a very important modification in yolov5 to achieve best accuracy in sewer inlet detection performance in UAV imagery, where YOLOv5x6 achieved Precision, Recall=90%, and 92%, YOLOv5x achieved Precision, Recall=78%, and 72%.Both P, R, and mAP are substantially better than the last results for the sewer inlet and manhole cover detection.

Now, urban water practitioners can create and update their inventory, the value added by the YOLOv5 detector is more than the incremental improvement that is usually gained by tuning the image classification method. Thus, this (S.I) detection solution can be used to develop the inventory of drainage infrastructure in urban areas.

7 References

[1]. Vitry, M.M.d., et al., Sewer inlet localization in UAV image clouds: Improving performance with multiview detection. Remote Sensing, 2018. **10**(5): p. 706.

[2]. Pasquet, J., Detection of Manhole Covers in High-Resolution Aerial Images of Urban Areas by Combining Two Methods. IEEE, 2017.

[3]. Commandre, B., Automatic reconstruction of urban wastewater Automatic reconstruction of urban wastewater Automatic reconstruction of ur. HAL open science, 2017.

[4]. Zhou, B., Smartphone-based road manhole cover detection and classification. Automation in Construction, 2022.

[5]. Agüera-Vega, F., F. Carvajal-Ramírez, and P. Martínez-Carricondo, Accuracy of digital surface models and orthophotos derived from unmanned aerial vehicle photogrammetry. Journal of Surveying Engineering, 2017. **143**(2): p. 04016025.

[6]. Mirko, S., ASSESSING THE IMPACT OF THE NUMBER OF GCPS ON THE. balticsurveying, 2019.

[7]. Jiménez-Jiménez, S.I., Digital Terrain Models Generated with Low-Cost UAV. International Journal o f, 2021.

[8]. Liu, Y., K. Han, and W. Rasdorf, Assessment and Prediction of Impact of Flight Configuration Factors on UAS-Based Photogrammetric Survey Accuracy. Remote Sensing, 2022. **14**(16): p. 4119.

[9]. J.Redmon, J., You Only Look Once:. IEEE, 2016.

[10]. J. Redmon, A.F., YOLO9000: Better, Faster, Stronger. IEEE, 2017.

[11]. Qiu, Q., Real-time detection of cracks in tiled sidewalks using YOLO-based method. Automation in Construction, 2023.

[12]. Zhu, J., Pavement distress detection using convolutional neural networks with. Automation in Construction, 2022.

[13]. Puliti, S. and R. Astrup, Automatic detection of snow breakage at single tree level using YOLOv5 applied to UAV imagery. International Journal of Applied Earth Observation and Geoinformation, 2022. **112**: p. 102946.
[14]. Wang, C.-Y., YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object. 2022.
[15]. Metashape, Agisoft. 2021: https://www.agisoft.com/downloads/installer/.
[16]. Wang, C.-Y., et al. CSPNet: A new backbone that can enhance learning capability of CNN. in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops. 2020.
[17]. Xu, R., et al., A Forest Fire Detection System Based on Ensemble Learning. Forests, 2021. **12**(2): p. 217.
[18]. Liu, S., et al. Path aggregation network for instance segmentation. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.
[19]. Deng, J., Y. Lu, and V.C.S. Lee, Concrete crack detection with handwriting script interferences using faster region-based convolutional neural network. Computer-Aided Civil and Infrastructure Engineering, 2020. **35**(4): p. 373-388.
[20]. Pearse, G.D., et al., Detecting and mapping tree seedlings in UAV imagery using convolutional neural networks and field-verified data. ISPRS Journal of Photogrammetry and Remote Sensing, 2020. **168**: p. 156-169.
[21]. Nie, M. and C. Wang. Pavement Crack Detection based on yolo v3. in 2019 2nd international conference on safety produce informatization (IICSPI). 2019. IEEE.
[22]. Wang, Q., et al. Improvements of YoloV3 for road damage detection. in Journal of Physics: Conference Series. 2021. IOP Publishing.
[23]. Zhu, J., et al., Pavement distress detection using convolutional neural networks with images captured via UAV. Automation in Construction, 2022. **133**: p. 103991.
[24]. Rydén, J. and S.E. Alm, The effect of interaction and rounding error in two-way ANOVA: example of impact on testing for normality. Journal of Applied Statistics, 2010. **37**(10): p. 1695-1701.
[25]. Babu, R.M.H., et al., Data science: a survey on the statistical analysis of the latest outbreak of the 2019 pandemic novel coronavirus disease (COVID-19) using ANOVA, in Data Science for COVID-19. 2022, Elsevier. p. 113-139.
[26]. Fawaz, E., Statistics lectures : ANOVA-test. 2022: https://www.youtube.com/@essamfawaz2001/videos/.
[27]. Niedoba, T. and P. Pięta, Applications of ANOVA in mineral processing. Mining Science, 2016. **23**.
[28]. Sow, M.T., Using ANOVA to examine the relationship between safety & security and human development. Journal of International Business and Economics, 2014. **2**(4): p. 101-106.