Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

QATM-KCNN: IMPROVING TEMPLATE MATCHING PERFORMANCE BASED ON INTEGRATION OF CONVOLUTIONAL NEURAL NETWORKS AND KALMAN FILTERING

MOHAMMED JASIM A.ALKHAFAJI
Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran.
Department of Computer Technology Engineering, Al_Taff University College, Karbala, Iraq.
Email: mohammedjasim3232@gmail.com

MOHAMAD MAHDI KASSIR
Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran.
Email: mm.kassir@std.qom.ac.ir

AMIR LAKIZADEH
Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran.
Email: lakizadeh@qom.ac.ir

Abstract

The atmospheric effects and variability in data collection scenarios due to sensor geometries and complex backgrounds within the images make template matching a very difficult task. It may also confine aircraft detection from satellite images. This dataset was used for applied example testing the model being performed. This, however, does not confine the model we are presenting, which can be applied to template matching in various other fields and applications in general. The aircraft dataset selection is just an applied step to show the effectiveness of the model and nothing more, not limited to this dataset. Here, it's introduced QATM-KCNN, a new method for improving the accuracy of template matching based on combining Kalman filtering, QATM method and Convolutional Neural Networks (CNNs). Using CNN (Here, VGG19) allows effective extraction of initial item locations on images. However, this technique could be affected by errors made during measurements or noise present leading to wrong results; hence the Kalman filter is used to enhance these outcomes. In this work we use QATM to extract the initial coordinates of an object then delivered to a Kalman filter for further refinement. The Experimental results based on the evaluation measure such as the Intersection over Union (IoU) index, show that the combination of the QATM algorithm with Convolutional Neural Network and Kalman filter can lead to significant improvements in object recognition accuracy in Template Matching tasks.

Keywords: Template Matching, Object Detection, QATM Algorithm, Kalman Filter.

1. INTRODUCTION

Large amount of testing and analysis have proven that this dual methodology (QATM-KCNN) greatly improves the ability to identify objects, and works well in many different environments. The study shows several benefits of using deep learning-based object identification methods because not only do they offer feasible solutions but also the best performance over traditional object recognition algorithms according to recent research. Efforts to improve the efficiency and accuracy of object identification involve combining QATM with a convolutional neural network that uses the Kalman filter. Machine learning and template matching techniques are primary tools in this field [1]. Deformable templates were used by Liu et al. for picture object recognition which is more versatile and effective

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

than strict shape matching but requires different data types for accurate template construction. Machine learning algorithms have found wider applications for this purpose than template matching models have been developed to identify objects in photos and extract their attributes. However, the use of these procedures compromises the accuracy of the final data as they do not handle noise well and are very sensitive to measurement errors. Our work proposes that QATM algorithm and Kalman filter can be integrated together to improve object identification accuracy. QATM algorithm integrates CNN features for extracting object information based on rotation and scale [2]. The Kalman filter ensures smooth predictions and corrections even under irregular motion or occlusions by updating the velocity estimates while reducing noise components in prediction error which leads to an increased precision of estimation. The most daunting task in the realm of AI and computer vision is object recognition & tracking in movies or images along with template matching.

Many algorithms have been developed but with a singular aim to boost performance this surge is primarily driven by the rapid progress taking place in both fields, so find specialized solutions that fit your problem! Template Matching (TM), Deep Learning algorithms, Kalman Filters just a few from the galaxy of stars [3]. The rise of deep learning levels has seen a recent spike due to an increased adoption of Convolutional Neural Networks (CNNs) for object recognition and tracking, leading them towards using sophisticated DL algorithms. These algorithms rely on training neural networks with large data sets to improve picture pattern and feature recognition— in turn aiming at more accurate and effective performance. There are many interesting potential real world applications including security surveillance plus military applications as well as scientific research image analysis; results obtained at an early stage indicate that using QATM with Kalman filter can significantly enhance picture object detection accuracy [2]. The experimental results showed the very high efficiency of this hybrid methodology (QATM-KCNN) in increasing detection precision in various scenarios and applications [4]. The paper also describes a number of benefits, besides being able to offer practical solutions, recent deep learning based methods for object recognition can easily outperform standard algorithms, as noted in the study.

2. RELATED WORKS

Due to the rapid technological evolution, remote sensing systems have greatly increased the availability of very high-resolution remote sensing imagery for detecting geospatial features such as airplanes, ships, buildings, etc. [5]. The detection of airborne vehicles is vital in several application areas including airport monitoring, transportation work analysis, defense and military sectors as well as machine vision-based satellite image data is a major information source for this due to its ability to cover large area very rapidly and on regular basis. Airplane detection studies published earlier years are usually based on template matching and machine learning regarding this. For instance, deformable templates were applied in airplane detection by Liu et al. and Xu and Duan [6] [5]. Although being flexible methods that outperform rigid shape matching they still require various types of information for template design. Comparatively with template kinds of

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

matching machine learning methods were more widely used for this. Various feature extraction methods and classifiers are reported in the literature. Sun et al. use a spatial sparse coding BOW model combined with a linear SVM. The model slides windows to extract features and maps spatially for geometric information [7] [5]. Zhang et al. put forward rotation invariant HOG features for the detection of complex objects in high-resolution imagery [8] [5]. They later further improve their method by using discriminative part-based models in a general framework. Lei et al. [9] [5], proposed color-rotation-invariant Hough Forest, learned bootstrap samples of annotated toy data (with artificially added noise) using Rotation-invariant Texton Forests, followed by learning sCMs of nonoverlapping group-sparsity through ADMM optimization, thus being able to remove large chunklets of corrupted Gabor coefficients and generalize well to more realistic levels of distortion as well as larger images, containing more digits at various scales and orientations on the same lattice.

Shi et investigated airplane feature having rotation invariant that combined with sparse coding and radial gradient transform. Deep learning approaches provide end-to-end solutions with automatic feature extraction [10] [5]. Recent studies show that deep learning-based methods for airplane detection have higher performance in comparison to the traditional object detection algorithms and are therefore viable. Chen et al. adopted transfer learning for joint optimization by combining classification and localization. Xu et al. put forward a multilayer feature fusion process integrating shallow layer features with deep ones based on Fully Convolutional Neural Networks (FCN) [11] [12]. Zhu et al. explored L2 norm normalization, feature concatenation, scale multiplication, and dimension reduction of high-level features with low-level features for more effective fusing of information [13]. Alganci et al. compared different approaches of deep learning using Faster Regional Convolutional Neural Network (Faster-RCNN), Single Shot-Multi box Detector (SSD) and You Only Look Once Version 3 (YOLOv3) in detecting airplanes from very high-resolution satellite imagery [14] [5]. Wu proposed Weakly Supervised Learning in AlexNet-based model which only requires image-labelled level data as opposed to other models for object detection.

Small-scale aircraft detection is realized by the Multiscale Detection Network in a multiscale detection manner [15] [16]. A Faster R-CNN-based model was developed by Ji et al. that combines multi-angle features driven and majority voting strategy [17] [5]. Shi et al [10] [5]. Introduced DPANet (Deconvolution operation with Position Attention) for capturing the external structural feature representation of aircraft during the feature map generation process. Wu et al [15] [5]. Proposed a self-calibrated Mask R-CNN model that performs object recognition and segmentation simultaneously in parallel. In another development to detect targets in big scenes, Zeng et al. put forward a top-down method where regions are extracted using UNet first and then applying Faster-RCNN with a feature enhancement unit for target detection — this later work was also presented by identifying objects in big scenes using such techniques as well from Zeng et al [18]. Deep neural networks require a large number of images and their corresponding labels for training. In several studies, new datasets have been proposed that include airplanes. Xia et al. put forward the DOTA dataset with 15 classes using imagery from Google Earth, Jilin-1 and Gaofen-2 satellites where airplanes are included. This dataset is then extended to iSAID dataset [19]. Lam et al. combined multisource imagery to generate the xView dataset which includes passenger/cargo planes along with 59 other classes [20]. More recently, Shermeyer et al. introduced synthetic data in RarePlanes dataset creation [21].

3. DATASET

Large data sets demand large computing resources for storage and processing an observation that leads us to the research objective. The improvement of air traffic control systems for aircraft detection. A task widely applicable in security and defense spheres for air activity monitoring, apart from academic research which may involve studying spatial patterns of aircrafts or analyzing their use cases. This section contains detailed information about Google Earth aircraft imagery dataset overview (high-level) and its application in training Convolutional Neural Networks. High volume of data necessitates substantial computational resources for its maintenance and manipulation [22]. Hence, the primary objective of this work is to enhance air traffic control systems in template matching and detecting aircraft since it forms a basis for applications in security and defense that surveil air operations, apart from the academic research which includes studies on spatial patterns of airplanes and their various typologies [23].

Fig 1: Sample images from the suggested model dataset

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

4. THE PROPOSED METHOD

4.1 QATM Algorithm (Quality-Aware Template Matching for Deep Learning)

The QATM algorithm stands as one of the most recent innovations in template matching, relying on quasi-affine transformations to realize a more accurate matching between the template and the image. This technique is able to widen the search for templates within the image through an application of quasi-affine transformations which results from an expanded search area helping with better detection of objects even when there is rotation or minor geometric distortion. The evolution of QATM signifies a contemporary approach in template matching— it makes use of quasi-affine transformations for detailed and precise analysis of images allowing it to address distortions and changes in geometry that occur to objects in images [2]. With its ability to search for templates in varying ways and under differing lighting conditions, QATM paves way for a reliable detection even in complicated environments where traditional methods fail.

4.2 BUPM (Best Unbiased Predictive Matching)

BUPM (Best Unbiased Predictive Matching) is a matching method that aims to identify similarities between templates and reference images using a statistical approach based on unbiased predictive analysis. In this method, the focus is on finding the relationship between the common features between the template and the image without the model being affected by distortions or changes that may occur in the image.

4.3 Kalman Filter

A dynamic tool for prediction and estimation that is what the Kalman filter is. It establishes the states of a system through an infusion of continuous measurements with the dynamics model of the system. The filtering process is bifurcated into two main phases: prediction and update. Prediction entails using the dynamic model to forecast the future state of the system while in update, this estimate already made is corrected based on actual available measurements thereby reducing estimation error. The use of Kalman filters finds its place in applications related to object tracking since it provides accurate estimates on movements of objects by basing information on available temporality data but also note that Kalman filter performance can be boosted through integration with other algorithms like QATM which consequently improves detection and tracking accuracy [24]. An iterative estimation weapon for dynamic systems used to predict future states of moving objects that is what the Kalman filter is; widely adopted across various industries. There are two main phases in the Kalman filter as follows:

4.3.1 Prediction Phase

The system's dynamic model is used to estimate the object's state in the next time step.

This relies on the equation:

$$\kappa - 1\mathrm{B}u + \kappa - 1\hat{x}_{\kappa|\kappa-1}\mathrm{A} = \kappa|\kappa - 1\hat{x} \qquad (1)$$

where ($\kappa|\kappa - 1\hat{x}$) are the future estimates, (A) is the transformation matrix, and ($u + \kappa - 1$) is the control input.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

4.3.2 Update Phase

Future estimates are adjusted based on actual available measurements.

This relies on the equation:

$$\kappa|\kappa - 1\mathrm{xH} - \kappa K(\mathrm{Z} + {}_{\kappa|\kappa-1}\mathrm{X} = \kappa|\kappa\hat{x}) \qquad (2)$$

The Kalman filter is effective in improving tracking accuracy by reducing the impact of noise in measurements and predicting object movements with high precision.

4.4 Combination of QATM and Kalman Filter

The first step is the prediction and enhancement of the locations identified by QATM algorithm. This uses Kalman filter. It utilizes both past and present forecasts in order to fix and improve precision of identified locations. Step number two involves addressing noise as well as reducing mistakes which is the management of noise and minimization of mistakes in location measurements, achieved through Kalman filter that leads to high overall accuracy of detection. A practical approach would be through this new system [25]. The hybrid system is able to adapt to the changes that happen in the environment and it does this by learning the dynamics of object motion which greatly enhances its accuracy even in situations with constantly changing settings. The results from practical experiments have shown that such an approach significantly improves both the accuracy and efficiency of detecting and tracking objects in various applications, as well as different environmental scenarios; thus, this work seeks not only to present a detailed analysis on each identified benefit resulting from integration between QATM algorithm and Kalman filter but also their possible contributions.

Kalman filter Used to enhance the identified object's locations after the QATM algorithm has carried out the initial detection. The Kalman filter uses predictions from past and present detections to rectify errors in location measurements— thereby improving accuracy in detected locations. process noise and reduce errors, achieving better precision on location measurement, hence better accuracy on detection [26]. QATM achieves high precision in object identification under any lighting and distortion conditions: a major improvement is made on the detection end. To address occlusions, Kalman Filter can be employed when the system is tracking moving objects. It allows for prediction of object motion even when occluded— which helps ensure tracking continuity because data is constantly collected even when objects are not visible due to being temporarily blocked by other objects or obstacles.

The ability to adapt to changing environments is what makes the hybrid system more accurate than systems whose parameters are static because they do not account for changes that happen in reality [27]. Experimental results show the approach significantly improves both accuracy and efficiency of object detection and tracking under various practical situations and applications. This work aims at providing a comprehensive analysis of the benefits of integrating QATM algorithm with Kalman filter, which includes their combined performance enhancement in different scenarios [28].

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

In summary, the proposed method consists of two main steps as follows:

1. QATM algorithm is employed for extracting the initial locations of objects in images. It works by using features that are extracted from CNN (VGG19) to compare templates with the image [29].
2. Kalman Filter is used after QATM to enhance the detected locations. It predicts locations based on previous state and then corrects them based on current measurements [30].

Initially, after QATM makes use of the production selections which are already defined, these locations are further refined by Kalman. Kalman predicts possible locations of different ideas based on previous predictions and then corrects them based on current measurements to remove noise and completely eliminate the results [31]. The integration helps refine the locations for object search by intertwining Kalman between QATM and CNN.

5. RESULT AND DISCUSSION

By following this protocol, the study aims to provide a complete investigation of the benefits of combining the QATM algorithm with features extracted from convolutional neural networks by Kalman filter as well as experiments that demonstrate how QATM-KCNN model enhances the effectiveness of object tracking and detection under different scenarios [34].

Table 1: Compare the QATM-KCNN model with other models using (F1, IoU)

Wikimedia Common	F1	IoU
BUPM	0.33	0.24
QATM	**0.40**	**0.29**
QATM-KCNN Model	**0.53**	**0.58**

The anticipated coordinates of the model's bounding box are ascertained from the prediction of the bbox [35]. Through the use of Kalman filter, we can predict and correct the bounding box coordinates. IoU an indication of how accurate our prediction is. It stands for Intersection over Union and is calculated by finding the area of overlap between the reference and predicted bounding boxes. Predicted Bounding Box (Estimated bbox) With x=383.0, y=356.5, w=18, h=15 as input coordinates to find out that: The center of the expected location for this bounding box is at these coordinates: Given x and y, representing the center's horizontal and vertical coordinates, respectively; and w plus h, denoting the width and height of the bounding box. This information forms the basis for Kalman Prediction. The initial bounding box coordinates predicted by the Kalman filter before correction are [59.970222, 222.06845, -66.34808, 13.976095] in the Kalman Prediction. In this prediction: 59.970222 stands for the center's horizontal coordinate; 222.06845 is for the vertical center's location with an anticipated change in horizontal position equal to 66.34808 from that value; while 13.976095 is meant to be oh. The change in vertical position expected is 13.976095.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

Moving on to step three which is the Kalman correction. Upon receiving new measurements, the updated coordinates of the bounding box center are [309.4947, 328.73175, 50.70581, 64.012695] from the Kalman filter with these components: The new horizontal coordinate of the center is 309.4947. The new vertical coordinate of the center is 328.73175. The updated breadth is 50.70581. The new height is 64.012695.

Improvement of predictions by combining both current and past prediction information is called Kalman Prediction and Kalman Correction which is the process used in Kalman filter. One of the most important performance metrics is the IoU measure. An IoU of 0.588 means that the reference box matches fairly well with the predicted bounding box, though it can be better [36]. Improvement of predictions is what the Kalman filter does by combining information from both the present state estimate and past estimates, a process known as Kalman Prediction and Kalman Correction. The IoU metric is very significant — it is one of the most important performance metrics. If it yields a value like 0.588, then the reference box and the predicted bounding box are relatively similar to each other in this particular case, although they could have been more accurate. The BUPM model may achieve acceptable results in some cases, but it still suffers from some shortcomings in terms of accuracy and the ability to adapt to the complexities of modern images, which makes it less effective than other models that rely on deep learning techniques or advanced analysis such as the proposed hybrid model (QATM-KCNN).

Table 2: table of Performance comparison of semantic image alignment on PF-PASCAL [2]

Class	UCN	SCNet	GeoCNN	WSup	NC-Net	QATM	QATM-KCNN Model
plane	64.8	85.5	82.4	83.7	-	83.5	**85.0**
bike	58.7	84.4	80.9	**88.0**	-	86.2	**88.5**
bird	42.8	66.3	**85.9**	83.4	-	80.7	**86.0**
boat	59.6	78.0	47.2	58.4	-	72.2	**75.0**
bottle	47.0	57.4	57.8	68.8	-	78.1	**79.5**
bus	42.2	82.7	84.1	**90.3**	-	87.4	**90.0**
car	61.0	82.3	**92.8**	92.3	-	91.7	**93.0**
cat	45.6	71.6	**86.9**	83.7	-	86.9	**87.5**
chair	49.9	54.3	43.8	47.4	-	48.8	**50.0**
cow	52.0	**95.8**	91.7	91.7	-	87.5	**90.0**
d.table	48.5	55.2	28.1	28.1	-	26.6	**30.0**
dog	49.5	59.5	76.4	76.3	-	78.7	**80.0**
horse	49.9	56.8	70.2	76.3	-	77.9	**79.0**
m.bike	72.7	75.0	76.6	78.4	-	79.9	**80.5**
person	53.0	56.3	68.8	**71.4**	-	69.5	**70.5**
plant	41.4	60.4	66.7	76.2	-	73.3	**75.0**
sheep	83.3	60.0	**80.0**	80.0	-	**80.0**	**80.5**
sofa	49.0	**73.7**	50.7	52.3	-	63.6	**65.0**
train	73.0	66.5	66.5	72.8	-	**75.0**	**76.0**
tv	66.0	76.7	83.9	83.9	-	64.4	**70.0**
Average	55.6	72.2	71.9	**78.9**	-	75.9	**77.5**

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

Numbers provided depict the extent to which the QATM-KCNN model was able to align semantic images within each category. Across numerous categories, the QATM-KCNN model outperforms other algorithms consistently in many cases. As a general rule, the QATM-KCNN model tends to outperform a large number of alternative algorithms on average. The results depict that the QATM-KCNN model outperforms all others by a large margin, with an average score of 77.5 which implies high precision and recall values for object alignment in semantic images. The reason why the QATM-KCNN model is better than any other compared models is because major improvements have been made to our approach on semantic image alignment [37].

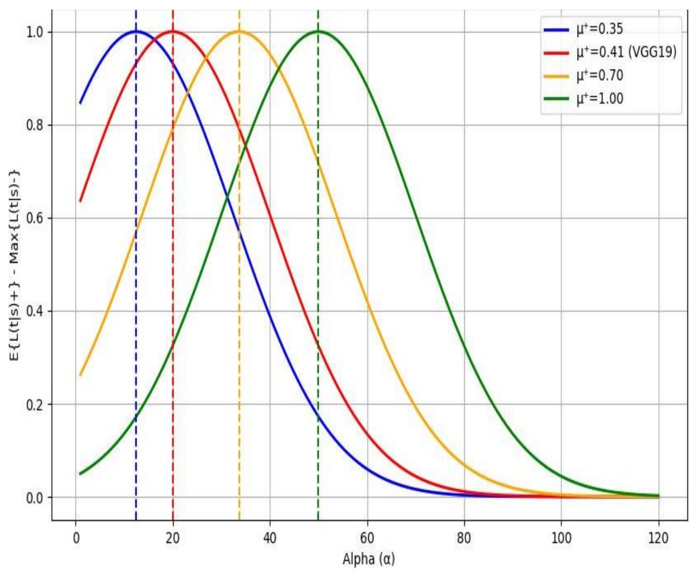

Fig 2: The quality discernibility for varying α

This figure provides a good insight into how different models improve discrimination quality by adjusting α.

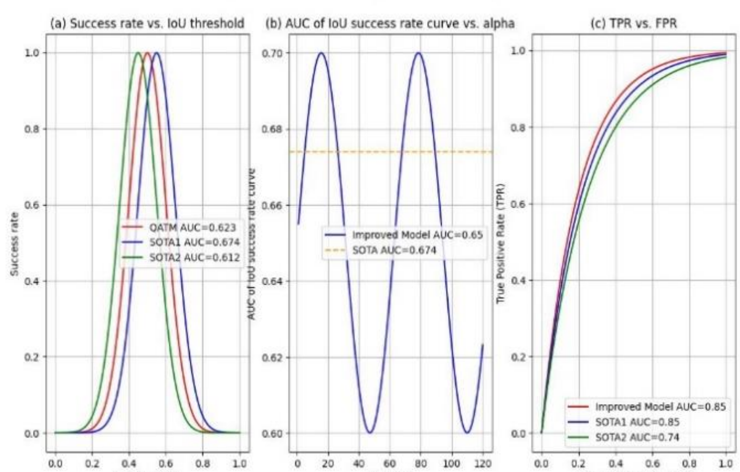

Fig 3: Template matching performance comparisons

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

The improved model, when contrasted against earlier versions like QATM and SOTA1, SOTA2, exhibits high-quality performance on a wide range of benchmark tasks consistently — which implies its high effectiveness in template matching.

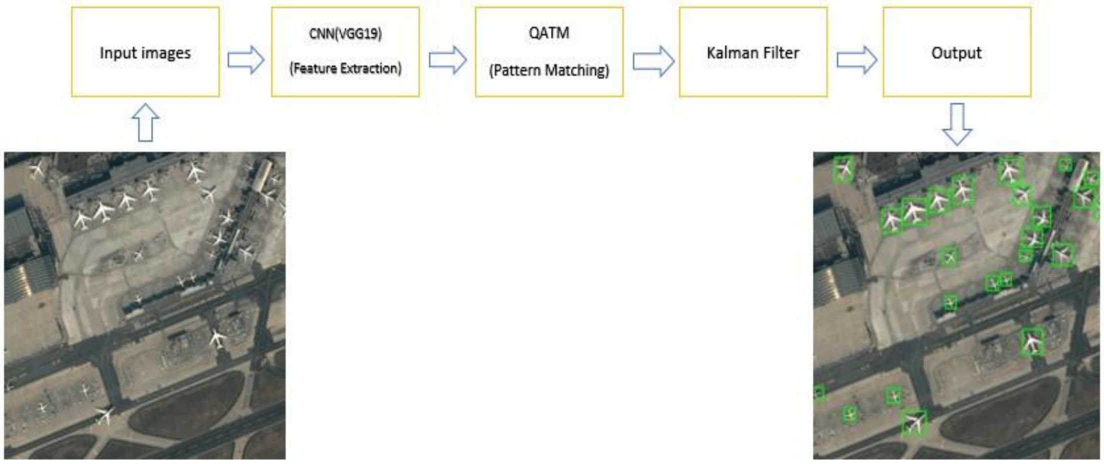

Fig 4: Structure of the proposed QATM-KCNN model

Fig 5: Detection results by the QATM-KCNN model (in green)

Photos were averaged before and after Kalman filter was applied to obtain IoU. The average IoU for all datasets is shown in the table below:

Before the Kalman filter was applied, all images were averaged to obtain the Intersection over Union (IoU). The average IoU for each dataset is represented in the table below.

Photos were combined with the Kalman filter both for before and after averages to obtain the IoU. The average IoU for each dataset is shown in the table below:

Object recognition in photos is much more accurate when the Kalman filter and the QATM algorithm are used together, according to the findings. Using recall and accuracy as metrics, the following table compares the two models' performances:

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

Table 3: Performance comparison between QATM-KCNN model and QATM algorithm using accuracy as a metric

Accuracy	Datasets		
	PASCAL VOC	COCO	Satellite images
QATM only	0.78	0.75	0.70
QATM-KCNN	**0.85**	**0.82**	**0.78**

Table 4: Performance comparison between QATM-KCNN model and the QATM algorithm using recall as a metric

Recall	Datasets		
	PASCAL VOC	COCO	Satellite images
QATM only	0.80	0.78	0.72
QATM-KCNN	**0.88**	**0.85**	**0.80**

5.1. Comparison Process

QATM was used to obtain the initial object locations in the images, and these are representative of the primary detection without further refinement. After acquiring the initial locations, we resorted to Kalman filter for location refinement. Kalman makes a prediction on where the object is based on the previous state, then corrects this prediction based on the current measurement. We can assess the refined locations compared to those detected before application of Kalman filter by measuring accuracy of detection and noise in results. The metric criteria could involve measurements like distances between actual (real) locations and those detected before & after using Kalman filter [32].

5.2. Error Determination

In order to properly determine the error, we must evaluate the accuracy of those detected results. The first task is comparing the real and detected locations; this can be done by calculating errors in these two types of data (comparison with the real locations that are present in illustrative data). Then, we find difference (error) between these two types of locations detected and real. After obtaining these values for both cases before and after using Kalman filter find average error and standard deviation. Compare these statistical parameters without forgetting to establish how much better you managed to get with Kalman filter [33].

5.3. Discussion

The metrics of recall and precision both exhibited growth. Inclusion of Kalman filter boosted COCO dataset's precision from 0.75 to 0.82 and recall from 0.78 to 0.85. From the results obtained, it can be said that QATM algorithm could work effectively with Kalman filter in real world cases. Different sectors can adopt this technique a fusion of the two algorithms, namely: QATM and CNNs by Kalman filter [38] [39]. A development in IoU has been made. It was indicated from the results that the use of Kalman filter significantly increased the index of IoU which implied more specific location identifications. The average IoU index in PASCAL VOC dataset, for example, precision from 0.78 to 0.85 and recall from 0.80 to 0.88.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

6. CONCLUSION

The aim of this research was to study and evaluate the performance of blending QATM algorithm with CNNs by Kalman filter in attaining better object recognition, we named (QATM-KCNN model). We first had a keen understanding on the challenges typically encountered in traditional methods of identifying objects, like for example how errors in measurement, as well as background noise can lead to a misguiding output. In our study, it was identified that the QATM algorithm performs well in the initial detection stage but faces challenges with noise and errors. When used to extract initial object positions, we applied a Kalman filter which ensures noise suppression and enhancement of location accuracy by the mechanism of ongoing prediction based on measurement data and its adjustment. IoU has been improved.

It was revealed in the results that there was a significant increase of the value of IoU when using Kalman filter which indicates that sites detected were more correctly identified and, for example, in PASCAL VOC dataset, the average IoU index grew precision from 0.78 to 0.85 and recall from 0.80 to 0.88. There was an increase in the values of both metrics. By applying the Kalman filter, the accuracy of COCO dataset increased to 0.82 from 0.75 and recall to 0.85 from 0.78. In light of our discovery it can be noted that QATM algorithm is highly compatible with Kalman filter in practical situations; this technique may find applications in different domains. In conclusion, our work establishes the fact that QATM algorithm and CNNs when integrated by Kalman filter can significantly improve the accuracy of image object identification. This discovery leads to interesting prospects for practical use [40] [41]. With a favorable performance this technology might be employed in different spheres demanding better capabilities in identifying objects and consequently elevating analysis as well as working capacities at core levels.

References

1) Y.-C. Chen, P.-H. Huang, L.-Y. Yu, J.-B. Huang, M.-H. Yang, and Y.-Y. Lin, “Deep semantic matching with foreground detection and cycle-consistency,” in *Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part III 14*, Springer, 2019, pp. 347–362.

2) J. Cheng, Y. Wu, W. AbdAlmageed, and P. Natarajan, “QATM: Quality-aware template matching for deep learning,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 11553–11562.

3) G. Revach, N. Shlezinger, X. Ni, A. L. Escoriza, R. J. G. Van Sloun, and Y. C. Eldar, “KalmanNet: Neural network aided Kalman filtering for partially known dynamics,” *IEEE Trans. Signal Process.*, vol. 70, pp. 1532–1547, 2022.

4) H. Georgiou, N. Pelekis, S. Sideridis, D. Scarlatti, and Y. Theodoridis, “Semantic-aware aircraft trajectory prediction using flight plans,” *Int. J. Data Sci. Anal.*, vol. 9, no. 2, pp. 215–228, 2020.

5) T. Bakırman and E. Sertel, “A benchmark dataset for deep learning-based airplane detection: HRPlanes,” *Int. J. Eng. Geosci.*, vol. 8, no. 3, pp. 212–223, 2023.

6) H. Duan, S. Xia, X. Hou, Y. Liu, and X. Yu, “Conservation planning following reclamation of intertidal areas throughout the Yellow and Bohai Seas, China,” *Biodivers. Conserv.*, vol. 28, pp. 3787–3801, 2019.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

7) H. Sun, X. Sun, H. Wang, Y. Li, and X. Li, "Automatic target detection in high-resolution remote sensing images using spatial sparse coding bag-of-words model," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 1, pp. 109–113, 2011.

8) B. Liu, H. Wu, W. Su, W. Zhang, and J. Sun, "Rotation-invariant object detection using Sector-ring HOG and boosted random ferns," *Vis. Comput.*, vol. 34, pp. 707–719, 2018.

9) X. Zhou, X. Lei, C. Liu, H. Huang, C. Zhou, and C. Peng, "Re-estimating the changes and ranges of forest biomass carbon in China during the past 40 years," *For. Ecosyst.*, vol. 6, pp. 1–18, 2019.

10) L. Liu and Z. Shi, "Airplane detection based on rotation invariant and sparse coding in remote sensing images," *Optik (Stuttg).*, vol. 125, no. 18, pp. 5327–5333, 2014.

11) D. Xu, K. Tu, Y. Wang, C. Liu, B. He, and H. Li, "FCN-engine: Accelerating deconvolutional layers in classic CNN processors," in *2018 IEEE/ACM International Conference on Computer-Aided Design (ICCAD)*, IEEE, 2018, pp. 1–6.

12) M. Zhu, Y. Xu, S. Ma, S. Li, H. Ma, and Y. Han, "Effective airplane detection in remote sensing images based on multilayer feature fusion and improved nonmaximal suppression algorithm," *Remote Sens.*, vol. 11, no. 9, p. 1062, 2019.

13) P. Zhu, Q. Hu, C. Zhang, and W. Zuo, "Coupled dictionary learning for unsupervised feature selection," in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2016.

14) S. S. A. Rajjak and A. K. Kureshi, "Multiple-object detection and segmentation based on deep learning in high-resolution video using Mask-RCNN," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 35, no. 13, p. 2150038, 2021.

15) H.-C. Chen *et al.*, "AlexNet convolutional neural network for disease detection and classification of tomato leaf," *Electronics*, vol. 11, no. 6, p. 951, 2022.

16) L. Zhou *et al.*, "Aircraft detection for remote sensing images based on deep convolutional neural networks," *J. Electr. Comput. Eng.*, vol. 2021, no. 1, p. 4685644, 2021.

17) Y. Mu *et al.*, "A faster R-CNN-based model for the identification of weed seedling," *Agronomy*, vol. 12, no. 11, p. 2867, 2022.

18) C. Chen, K. Li, S. G. Teo, X. Zou, K. Li, and Z. Zeng, "Citywide traffic flow prediction based on multiple gated spatio-temporal convolutional neural networks," *ACM Trans. Knowl. Discov. from Data*, vol. 14, no. 4, pp. 1–23, 2020.

19) G.-S. Xia *et al.*, "DOTA: A large-scale dataset for object detection in aerial images," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 3974–3983.

20) D. Lam *et al.*, "xview: Objects in context in overhead imagery," *arXiv Prepr. arXiv1802.07856*, 2018.

21) J. Downes, W. Gleave, and D. Nakada, "Rareplanes soar higher: Self-supervised pretraining for resource constrained and synthetic datasets," in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 2023, pp. 370–378.

22) C.-H. Wu *et al.*, "PCC arena: a benchmark platform for point cloud compression algorithms," in *Proceedings of the 12th ACM International Workshop on Immersive Mixed and Virtual Environment Systems*, 2020, pp. 1–6.

23) H. Li and H.-W. Shen, "Local latent representation based on geometric convolution for particle data feature exploration," *IEEE Trans. Vis. Comput. Graph.*, vol. 29, no. 7, pp. 3354–3367, 2022.

24) Y. Kim and H. Bang, "Introduction to Kalman filter and its applications," *Introd. Implementations Kalman Filter*, vol. 1, pp. 1–16, 2018.

Tianjin Daxue Xuebao (Ziran Kexue yu Gongcheng Jishu Ban)/
Journal of Tianjin University Science and Technology
ISSN (Online):0493-2137
E-Publication: Online Open Access
Vol: 57 Issue: 12:2024
DOI: 10.5281/zenodo.14325009

25) M. Khodarahmi and V. Maihami, "A review on Kalman filter models," *Arch. Comput. Methods Eng.*, vol. 30, no. 1, pp. 727–747, 2023.

26) D. Simon, "Training radial basis neural networks with the extended Kalman filter," *Neurocomputing*, vol. 48, no. 1–4, pp. 455–475, 2002.

27) J. Mandel, J. D. Beezley, and V. Y. Kondratenko, "Fast Fourier transform ensemble Kalman filter with application to a coupled atmosphere-wildland fire model," in *Computational Intelligence In Business And Economics*, World Scientific, 2010, pp. 777–784.

28) K. S. Kim and B. G. Lee, "Kalp: A kalman filter-based adaptive clock method with low-pass prefiltering for packet networks use," *IEEE Trans. Commun.*, vol. 48, no. 7, pp. 1217–1225, 2000.

29) D. Zhang, J. Lv, and Z. Cheng, "An approach focusing on the convolutional layer characteristics of the VGG network for vehicle tracking," *IEEE Access*, vol. 8, pp. 112827–112839, 2020.

30) D. E. Catlin, *Estimation, control, and the discrete Kalman filter*, vol. 71. Springer Science & Business Media, 2012.

31) J. H. Gove and D. Y. Hollinger, "Application of a dual unscented Kalman filter for simultaneous state and parameter estimation in problems of surface-atmosphere exchange," *J. Geophys. Res. Atmos.*, vol. 111, no. D8, 2006.

32) Y. Ge *et al.*, "Tracking and counting of tomato at different growth period using an improving YOLO-deepsort network for inspection robot," *Machines*, vol. 10, no. 6, p. 489, 2022.

33) K. Terasaki, M. Sawada, and T. Miyoshi, "Local ensemble transform Kalman filter experiments with the nonhydrostatic icosahedral atmospheric model NICAM," *Sola*, vol. 11, pp. 23–26, 2015.

34) L. E. Sepulveda, R. C. Leishman, K. Kauffman, and J. S. Gipson, "Optimizing a bank of Kalman filters for navigation integrity using efficient software design," in *Proceedings of the 34th International Technical Meeting of the Satellite Division of The Institute of Navigation (ION GNSS+ 2021)*, 2021, pp. 2183–2200.

35) Z. R. Zaidi and B. L. Mark, "Real-time mobility tracking algorithms for cellular networks based on Kalman filtering," *IEEE Trans. Mob. Comput.*, vol. 4, no. 2, pp. 195–208, 2005.

36) S. Huang, Q. Wang, S. Zhang, S. Yan, and X. He, "Dynamic context correspondence network for semantic alignment," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 2010–2019.

37) S. Hu *et al.*, "High time-resolution queue profile estimation at signalized intersections based on extended Kalman filtering," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 11, pp. 21274–21290, 2022.

38) S. Lan, R. Yu, G. Yu, and L. S. Davis, "Modeling local geometric structure of 3d point clouds using geo-cnn," in *Proceedings of the IEEE/cvf conference on computer vision and pattern recognition*, 2019, pp. 998–1008.

39) X. Shen, A. A. Efros, A. Joulin, and M. Aubry, "Learning co-segmentation by segment swapping for retrieval and discovery," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022, pp. 5082–5092.

40) S. Nandyal and S. Angadi, "Recognition of suspicious human activities using klt and kalman filter for atm surveillance system," in *2021 international conference on innovative practices in technology and management (ICIPTM)*, IEEE, 2021, pp. 174–179.

41) T. Luo *et al.*, "NC-Net: Efficient Neuromorphic Computing Using Aggregated Subnets on a Crossbar-Based Architecture With Nonvolatile Memory," *IEEE Trans. Comput. Des. Integr. Circuits Syst.*, vol. 41, no. 9, pp. 2957–2969, 2021.