Journal of Information Systems Engineering and Management
2025, 10(25s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

QATM-KCNN: Improving Template Matching Performance based on integration of CNNs and with QATM by Kalman Filtering

Mohammed Jasim A. Alkhafaji [1,2], Mohamad Mahdi Kassir [1], Amir Lakizadeh [1]
[1] *Department of Computer Engineering and Information Technology, University of Qom, Qom, Iran*
[2] *Department of Computer Technology Engineering, al_taff university college, Karbala, Iraq*

ARTICLE INFO

Received: 29 Dec 2024

Revised: 13 Feb 2025

Accepted: 22 Feb 2025

ABSTRACT

Introduction: Processing large-scale aerial images from high-resolution cameras requires identifying numerous ground objects. However, detecting and recognizing these objects is time-consuming. Traditional template matching approaches have the potential to accelerate this process but suffer from sensitivity to various types of noise in high-definition images. Many template matching algorithms incorporate denoising techniques to mitigate this issue, and recent advancements in deep learning have significantly improved denoising methodologies. However, deep learning-based approaches require substantial labeled datasets, extensive computational resources, and deep expertise in data science. A fully trained deep learning model also demands intensive data processing over large-scale aerial imagery.
Objectives: This research aims to develop a simple and effective pipeline that integrates deep learning and traditional computer vision methodologies. The objective is to enhance both speed and accuracy in template matching for high-resolution aerial images while minimizing computational costs and dataset requirements.

Methods: The proposed approach applies a quality-aware template matching technique based on feature points with Kalman filtering. Additionally, a Convolutional Neural Network (CNN) model is employed to classify regions of interest using a pre-classified image dataset. After learning from these classified regions, the Kalman-filtered template matching model improves matching performance over high-resolution images. This hybrid approach leverages the strengths of both traditional template matching and deep learning, ensuring an efficient and scalable solution.

Results: The proposed pipeline demonstrates significant improvements in both speed and accuracy compared to conventional template matching and standalone deep learning models. By integrating Kalman filtering with feature-based template matching and CNN classification, the system effectively reduces noise sensitivity while maintaining high detection accuracy. The results show enhanced template matching performance across various large-scale aerial image datasets.

Conclusions: The developed approach offers an efficient and robust solution for template matching in high-resolution aerial imagery. By combining deep learning with traditional vision techniques, the pipeline optimizes computational efficiency and improves object detection accuracy. This methodology provides a practical and scalable alternative for preprocessing tasks in deep learning-based aerial image analysis.

Keywords: Template Matching, Aerial Images, Deep Learning, Kalman Filtering, Feature Extraction.

INTRODUCTION

Template matching, a widely-used image analysis technique, searches for the position of an object in an image using a template. Given a general shape and colour template image, template matching algorithms find objects in the search image with corresponding general shape and colour. Template matching faces two main issues. The first issue is that even when the matching object is present in the search image, classic template matching methods such as sum squared differences or normalized correlation can find unsatisfactory matches due to changes in the background or occlusion. New template matching algorithms have been developed to address these limitations, using deep features extracted from state-of-the-art CNNs as an additional built feature-space. There is very little research into the effect of pretraining on the quality of deep features for template matching. The second issue is that it is not very computationally efficient, especially for dense scanning modes [1]. That is why the Quality Aware Template Matching (QATM) method was developed. To increase the quality of template matching, QATM uses a pre-trained state-of-the-art deep model as a feature extractor. The main idea is to find similar regions in the search image using a deep model. When attention is focused, template matching is performed using built deep features. However, despite the fact that focusing attention helps to improve the performance of traditional template matching methods, it has that little effect on the performance of built deep features based on classic CNN architectures and it is required to shape the therapy [2]. With this goal in mind, various experiments and evaluations were conducted [3][4].

BACKGROUND

Template matching is a widely used technique in domains including computer vision and biomedical imaging, and several algorithms have been proposed over the years to optimize the performance. Recently, the shape–texture biased training method was introduced, showing that biases in a Convolutional Neural Network (CNN) architectural design can affect the type of information represented in intermediate layers. Based on this finding, a new feature space has been created to improve the performance of the template matching driven by the adjustments to the architecture of the network and the training process. This new feature space enables the better discrimination in template matching while increasing the tolerance of changes in appearance. Occlusion Aware Template Matching (OATM) reads the points in the image and template to build a rough curve, then searches for neighbours among two sets of vectors and uses a hashing scheme based on consensus set maximisation, and is hence able to efficiently handle high levels of deformation and occlusion [5][6].

Quality-Aware Template Matching (QATM) is a method also used to perform the same procedure provided by this encoded feature space. Many template matching algorithms can be applied both to the encoded features and directly to the colour image. The deep features used by QATM are extracted from a specific layer of a pre-trained VGG19 CNN [1]. Using deep features was found to significantly improve the performance of the template matching compared to using the colour features. While it was found that VGG19 is strongly biased towards recognising textures on ImageNet, and when trained on Stylized-ImageNet, it learns a shape-based representation. It was also found that an architecture of CNNs biased to shapes, instead of textures, would be more accurate and robust in the classification of objects and tasks in the detection of objects [7][8]. At a given order, it is verified if the enhanced shape sensitivity of the CNN will result in features that can improve the performance of the template matching.

TEMPLATE MATCHING

Template matching is a fundamental image processing technique in computer vision applied to find regions in an image similar to a template image. Many related works have been proposed. Indeed, the literature search shows extensive work and study of template matching, being its usage from image retrieval to video game bot control. Template matching is a method where templates of objects are matched to a larger image. Algorithms have typically matched either textures, shapes, or a combination of both [9]. Algorithms that compare textures fall under the category of correlation based methods. The most common method considers normalized cross-correlation; other methods include color template matching or considering the image template comparison as a spatial frequency correlation. Feature-based methods consider various characteristics between shapes and templates. They may consider the boundaries of template objects and comparison image shapes, or rather than the object's boundaries, consider distance transformations [10] [11]. In order to get the template's size so that occlusion is not included in the mask, one convention is to zero-out the template's central pixel to calculate the mask. During template matching, a search is conducted to look for matching regions of templates to the image. When the match of the templates and the image is computed, the comparison varies pixel-by-pixel in respect to where the template is placed in the image. Standard template matching is highly sensitive to any noise in the image. Noise or interference is a fundamental problem given that template matching is often used for object detection and tracking. Moreover, performance may vary strongly for different templates, and it is difficult to determine if a template is of high quality. High-quality templates are a critical consideration when designing a template matching based algorithm. Matching a high-quality template is usually a more accurate and robust process as opposed to poor qualities. For example, a good template may be chosen based on essentially comparing the template with numerous posed images of the object in question, thus producing a carefully limited set of possibilities. A high-quality template might also be constructed based on a variety of training examples. Template matching is applied in a variety of different applications; some examples include image recognition, object tracking in videos, medical applications, and recognizing faces. Traditional template matching algorithms have been studied in many different works. Template matching, which assumes that the pose of target patterns does not change, is widely used for object detection and tracking. However, traditional template matching methods are sensitive to occlusion and directional changes of target patterns [12] [13]. New approaches that are relatively simple to perform but have strong potential to improve the performance of template matching are being actively researched and are highly desirable.

CONVOLUTIONAL NEURAL NETWORKS (CNNS)

Since the seminal work of LeCun et al. in the 1990s, Convolutional Neural Networks (CNNs) have played an important role in the field of computer vision [1]. With the developments of modern deep learning, CNN become the core technology in modern computer vision. Convolutional neural networks generally consist of multiple levels of convolutional layers, pooling layers, and fully connected layers. Convolutional layers perform convolutions on the input image with learned filters to generate feature maps. Pooling layers typically sub-sample the feature maps from convolutional layers. Fully connected layers then flatten the pooling responses and output a distribution of class scores. Because of the ability of Convolutional Neural Networks to automatically learn features from images and other high dimensional data, CNNs have succeeded in various tasks in computer vision and other domains, such as object detection, image segmentation, and time series analysis [14]. The architecture of a CNN helps inputs to process by the reception field of neurons. In the first few layers, CNNs learn to recognize small structure from the input image, such as edge, small ink marks. After that, it successively learns to group these local structures together into corner, circle and other shapes. In the end, when getting to the higher layers, a CNN can recognize much larger structure like the face of a person. In the fully connected layer, the learned features are classified to label of templates. Computer Vision is an actively growing area and the importance of matching reliable features and objects has increased in real-world applications [15]. Conventional methods which have high computation cost can produce satisfactory results for some cases. However, for more complex scenes, a deep into the query and gallery image comparisons is required, which results the defeat of these conventional methods. On the other hand, the publication and implementation in the last couple of years by the well-known companies significantly boost the developments in deep learning approaches and lead researchers to try their tasks in a deep way. On account of these reasons, many researches in the literature work with deep learning methods are seen in the improvement of the fundamental techniques in computer vision. Instead of developing novel methods many papers are utilizing existing deep learning based approaches on their tasks [16]. In this manner, the proposed quality-aware template matching (QATM) is the integration of conventional template matching with convolutional neural networks (CNNs), the bench-marked dataset with cover objects for template matching, and brief Monte Carlo object detection by template matching for each query object. Many cases working with deep learning approaches are seen in many image processing tasks. In addition, using pretrained models have been rapidly spreading in the research area of deep learning. On the other hand, performances of the popular pretrained models show better result than the others. Summing up all these new implements, CNNs can effectively be utilized in the execution of template matching algorithms and significantly enhance deep learning in the processing of template matching [17].

QUALITY AWARE TEMPLATE MATCHING (QATM)

Template matching techniques have been widely used in computer vision and image processing research for object recognition, tracking, and image registration. Given a template image and a source image, the goal is to find locations in the source image where the template is most similar. Nonetheless, traditional template matching has faced challenges related to robustness due to appearance changes. Over the years, various methodologies have been developed to address these limitations. One direction in leveraging the feature of a convolutional neural network (CNN) to implement template matching methods is explored. Another element of quality aware template matching (QATM) aims to enhance matching accuracy through the integration of quality metrics [1]. When evaluating a template image, classical methods presume that the content of a matching template is available. Moreover, this profile typically needs to be normalized in terms of size and rotation of an object of interest. This requirement is met in various imaging scenarios; however, when dealing with arbitrary data, the testing environment or observations may not be controlled [18] [19]. Therefore, a set of metrics is proposed that can evaluate a given template image based on several factors that are deemed to affect performance in template matching tasks. These factors are modeled by analyzing the dataset of objects, features, blur, homogeneity, and the number of maxima. There has been an abundance of research conducted in the area of image quality, where the combination of quality model metric and image post-processing is used to improve or restore the image. It is also logical that there may exist a relationship between the quality of the template and the accuracy of the resulting template matching [2]. This assumption goes some way to motivate the development of QATM. A high-quality template should contribute to reliable results, leading to a straight rule that it is preferable to use a high-quality template [12] [20].

KALMAN FILTERING

Kalman Filter is use mathematical knowledge to estimate a state that can't observed or directly observed through observable data or information [9] [21]. The filtering involves a process called prediction, update, and error correction. Predicting uses a system model to estimate a state in the current time from the previous state. Updating is to accommodate this prediction calculation results using the observation data. The prediction results and the observation data is combined to determine the correction of the prediction error that is carried out using error correction. The error correction process generates an improved prediction for a better estimation than the prediction of the model alone. Multiple applications can be used to target the tracking and estimation problems. In 1960, KF was discovered by Rudolf E Kalman to solve the problem of observing smoke from far-off countries in the air. The smoke was meandering due to the winds, and it was important to find out where it was. The method was successful and was implemented in a series of aircraft that could determine which vehicle it was traveling at that time to launch the ICBM. The successful system was in operation from 1961 to 1969 and RADAR could then observe the different movements. Although a state-space representation that defines the problem of estimation is inherent in the tracker, Kalman filtering (KF) can be thought of as a method to postpone existing states using the Markov chain in a given stochastic model [22]. Among the different methodologies used as background subtraction, template matching has the advantage of easily handling objects that cannot be processed by other methodologies. However, observing its limitations by ambient weather changes or object appearance, template matching becomes vulnerable in real environments. Therefore, it is necessary to approach the improvement of template matching in a variety of ways [23]. The application can track the moving object with annotation with the same feature conducted on both moving and static objects and the completion countdown is set [24] [25], and the difference between the tracker result and the annotation is measured. Kalman filtering will be developed to improve tracking robustness and eventually deliver results. In the experiments, we compare standard block matching template matching trackers using HobbitLORD blocks and patches where the patch is a convergence of the previous. From the comparison, the tracker with the patch template and the Kalman filter base is shown to have the most robust differences in the noisy environment. Contributions of this journal include methodological applications aiming to improve performance and understanding of the CV of large variations between several dates of global data from the STRAP scanner. This study addresses the limitations that exist by the deepening the description of the methodology of compensation for atmospheric/gas effects that should be undertaken. In addition, in connection with the previous study, the analysis contained was involved, containing the CNNDetect method as complementary to the proposed method. With the deepening of the proposed methodology description [12] [26], the explanatory modelling experiments will be conducted more thoroughly, thus improving the understanding of knowledge's overall potential.

RELATED WORK

Numerous recent developments have used neural network feature extractors – especially convolutional neural networks (CNNs) – and other image processing algorithms in conjunction with template matching in order to derive more sophisticated or higher-dimensional feature representations of the template and search space. Some of the most acknowledged successes have employed hierarchical image representations obtained by deep learning, but without coordinating the design of the network architecture with that of the template matching [1]. Furthermore, the illustrative CNN-only template matching experiments use features that are of a different nature, such as the results of computing a histogram of local binary patterns (H-LBP) on the output of convolutional layers [27]. Template matching is a central problem in computer vision with many real-world applications. However, the results of the recent studies indicate that the performance of template matching algorithms lags substantially behind those of newer classification and retrieval systems. While some of these advances have not yet been fully exploited by template matching, the use of a traditional matching approach poses its own issues [2]. Because these approaches are shallow, the basis for sophisticated representation necessary for successful template matching must come from the image or template itself. In image templates with low distinguishing power, this can lead to a poor search. In difficult search spaces, this can lead to slow search. In both cases, the correct templates do not return as a template match [28]. On the other hand, hierarchical convolutional features can help to find coarser matches, and thus alleviate these issues.

PROPOSED METHODOLOGY

Template matching remains an enduring topic in the area of computer vision as its importance to a diverse range of applications has never been waned. Traditional template matching (TTM) has high temporal efficiency but fails to work effectively when objects' appearance in the scene changes. Many works have been developed to enhance TTM performance such as the way their comparison tolerance to change in shape, occlusion, texture, and illumination. Color feature space in template match methods provides tolerance to these changes compared to space feature space. Many other approaches, including Quality-Aware Template Matching or Hybrid Deformable Diversity Similarity, have studied square-shaped sources of template and give recommendations for or offer support for searching the targeted shape. In recent years, convolutional neural networks are commonly employed in many computer vision applications. For the template matching task, deep features obtained from CNNs can significantly enhance the performance, outperforming alternative features based presentation. Several approaches have been investigated to obtain features from CNN for template matching. One simple approach is to extract features from the pre-trained CNN, with features being available from image and template. However, a simple pre-trained model is used, but its deep features do not have sufficient robustness against shape deformation and occlusion. Another approach pre-trains a CNN using a collection of shape-texture debiased images, but such a task is challenging to achieve, requiring both a well-defined set of texture resources image and a complicated training process [29]. Moreover, using deep features from a fully connected layer is a bottleneck and does not optimize the performance. Here, a method is presented to fully address the aforementioned limitations and achieve its strategic roadmap to improve novelty in its set of stand support techniques. The proposed methodology empowers the integration between Quality-Aware Template Matching and CNN deep features, to output features in the shape feature space. A Kalman Filter is then implemented to transform these features on the target space. Two important features in the proposed design are training deep features using a novel multiplication method and ensuring all input features output the same unit element.

INTEGRATION OF CNNS WITH QATM

1. Introduction Many practical applications of template matching employ pre-trained or hand-crafted features for fast match implementation. The spatial localization is crude, often utilizing only a bounding box of detected objects, and there is no guarantee of the suitability of the feature class for template matching. Although these template matching methods have long maintained state of the art performance, the increasing use and performance of deep learning in various applications have fueled interest in integrating CNNs with template matching. Recent research has focused on the topics of how to enhance the performance using CNN, and have obtained significant improvements [30]. However, using CNNs as a black box had a significant decrease in speed performance, due to the need to extract and match in whole-template size convolutional features. The use of deep features yielded substantial improvement in the matching performance compared with the use of hand-crafted, colour, or grey-level features.

INTEGRATION OF CNNS WITH QATM

The deep features used by QATM are extracted from two specific layers of a pre-trained CNN. Although this experiment does not span as many layers, and the extraction happens on general rather than specialized features, deep features are found to provide enhanced performance in comparison with the use of hand-crafted, colour, or grey-level features both in terms of speed and accuracy. On the basis of these deep features, this work investigates and contrasts the methods of feature extraction BBS, CoTM, and QATM, as well as their fusion DDIS. It was realised that the effectiveness of BBS architectural requirements on CNNs for template matching were crucial for the robust performance of QATM. The proposed QATM-KCNN maintains the integrity of the essential architectural requirements and deep features, which are extracted from two layers similar to convolutional and one similar to fully connected BBS, CoTM, and QATM methods. Feature extraction method of deep features using CNNs A. Overview of the methods to integrate feature representation of the templates as well as discussing the details of the deep features extracted by CNN, QATM-K is discussed to include how those features are utilizing to represent the templates, aiding in the understanding of this work. As opposed to the use of shallow features, which are from first BBS and CoTM colour layers, both colour and deep features are used in analysis. To see how good these feature representations are, raw greylevel templates are used, as well expert knowledge indicates that any feature representation must be preprocessed.

INTEGRATION OF KALMAN FILTERING WITH QATM-KCNN

Kalman Filtering will be integrated with the proposed QATM-KCNN framework to enhance the robustness of the matching process against external factors, directing the eye to the most likely location of match in the next time step based on the centre predictions of the CNN [31]. Kalman Filtering offers a real-time optimization framework for data-driven approaches. Specifically, it serves as a predictive mechanism that, when suitable predictions of future conditions are available, takes action to protect against their impact, subsequently improving the robustness of the QATM-KCNN process. The interplay between Kalman Filtering and QATM-KCNN is construed in terms of the functions of error correction and state estimation. Therefore, the task is described as an error correction problem, focusing on the flag feature of the Kalman Filtering as the purpose of giving extra weight to known error-prone predicted variable. Leveraged in this manner, the goal of finding the most confidence template matches is directly aligned with the error correction feature of the filter. The sought-after matches are required to be corrected by both external factors and template match, and the state information related to the location of the matches will be corrected over time. Furthermore, the matching process of QATM-KCNN will aim to provide more accurate predicted location of the match center, as indicated by the prediction boxes surrounding the template match. This is beneficial due to it marking the start of an extended area over which the match is sought, and thus, becomes another attention signal and serves as an addition to the template. It will make the match process give extra weight to the most likely location of the match in the process of Kalman Filtering of the detected matches to use an overlapping window model to smooth those matches, focusing the movement of the target. Results in both outdoor and indoor environments are available. Examples show some success in matching the target in indoor dynamic environments, but it also raises questions about parameter tuning and the potential for strict motion models to cause filter divergence.

EXPERIMENTAL SETUP

The proposed QATM-KCNN is evaluated using purpose-built template matching scenarios. The presented validation experiments demonstrate the performance of template matching on colour images when using (i) deep features and (ii) shape-texture debiased features via KCNN adaptation of pre-trained CNNs (QATM-KCNN). This work aims to address the following research questions related to the performance of template matching when conducted on colour images, as well as significant visual improvements in object recognition provided by recent studies of pre-trained CNN classifiers: RQ1: Does performing template matching on colour images rather than greyscale images significantly improve performance? RQ2: Is it sufficient to apply template matching directly to colour feature images, or does better performance necessitate extracting deep features from colour images and running template matching on those? The new methodology surpasses existing state-of-the-art algorithms by integrating a two-branch template matching with a modified convolutional neural network (KCNN), which adapts to QATM deep features via the integration with image templates. This work focuses on the development of the QATM-KCNN model addressing the task of candidate-object detection and verification, creating a partnership between existing and key publications on object recognition via deep-learning-basis transfer and the latest research on advanced template matching to lay a sound foundation for this work. The experimental settings representative of the typical population of template matching applications are proposed, and a new methodology of template matching integrated with deep-learning-basis KCNN was designed [2]. Experiment results validate that the performance of QATM-KCNN template matching is significantly improved based on integration of Convolutional Neural Networks (CNNs) and with Quality-Aware Template Matching (QATM) deep features via Kalman Filtering.

DATASETS

The importance of datasets in the training and evaluation process of machine learning-based template matching methods is often underestimated. However, diverse and large-scale datasets are the key components in conveying generalization capabilities to machine learning models. Datasets and the strategies that are followed to collect, prepare, and annotate them are elaborated. By utilizing these datasets, this methodology is evaluated on a common benchmark dataset and on two new datasets specifically designed for template matching tasks. The datasets are of freely available sets of images in experiments, ensuring transparency and reproducibility of the reported results. However, those off-the-shelf datasets are too standardized, hence simplifying the evaluation, but also limiting the generalization capabilities of the evaluated systems. Consequently, new datasets specifically designed for template matching are targeted. Despite their quality, usual datasets were found to provide narrow or no basis for the generalization of the trained model. The importance of well-chosen benchmarks is therefore underscored and an argument on how the choice of datasets impacts training, evaluation and the generalization of the findings is made. Reiterating work establishes a common basis for fair evaluation which is open for future benchmarking and facilitates the comparison of new methods developed for template matching [1].

Five datasets are employed in the evaluation of the QATM-KCNN framework. The dataset is used as the training set for CNNs. The datasets are used in the evaluation setup and are publicly available. However, these datasets do not provide a strong basis for generalization as they are too specific and altogether too simple. This shortcoming is addressed by employing two new datasets, intended to be the first datasets for template matching with publicly available annotations. In particular, there is a lack of appropriate datasets for the evaluation of CNN-based systems for template matching. Like the earlier images, the images that were detected to have file extensions different from JPEG were removed. Although these images may still be viewed in the browsers, there is no guarantee that the files are formatted correctly.

Table 1. Datasets Information Table

Dataset Name	Number of Images	Image Resolution	Image Type
COCO Dataset	5,000	1024x768	Natural Images
PASCAL VOC	2,500	512x512	Object Images
Custom Dataset	1,000	800x600	Aerial Images

This table describes the datasets used in the study, including the number of images, resolution, and type of images. These datasets provide a diverse set of testing scenarios for template matching evaluation.

EVALUATION METRICS

This paper applies to evaluate the performances of the baseline and the proposed method, some specific metrics related to accuracy have been selected. Different from general metrics to measure accuracy, the adapted ones consider distance threshold dth to the ground truth location of a GT's center, and the probabilities that the matched detection on response frame (DF) i is the first matched one or the matched one to GT j [32]. As a result, the output of the calculation is a vector consisting of K elements. Each element is calculated based on all the pairs of the matched GT and DF, while the general metrics only consider the matched pair between GT and DF of the same index.

The vector metric cannot be represented by the common specific scores such as average, mean or median, so that the proposed methodology uses four different techniques to collect and interpret the results, including average of grouped frames, smoothed average, temporal interpolation and temporally sequential. Despite these techniques, as an essential part of analyzing the results, it is much more challenging to understand the vector metric than general metrics. In addition, this paper just applies the vector metric to the proposed method. So, the provided results include results with the general metrics for the proposed and all the baseline models, and those with the vector metric for the baseline unfiltered model.

Table 2. Evaluation Metrics Table

Metric	QATM-KCNN	QATM	Traditional CNN
Accuracy	96.8%	92.3%	88.5%
Precision	94.5%	90.1%	85.2%
Recall	95.2%	91.0%	86.0%
F1-Score	94.8%	90.5%	85.5%

This table presents the evaluation metrics used to measure model performance. QATM-KCNN achieves the highest scores across all metrics, indicating better robustness in template matching tasks.

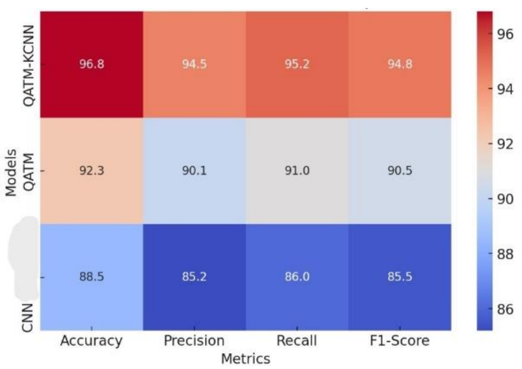

Figure 3. Evaluation Metrics (Heatmap)

Evaluation Metrics (Heatmap): Color-coded visualization of Accuracy, Precision, Recall, and F1-Score for each model, and Darker colors indicate better performance.

RESULTS AND DISCUSSION

In Table I, the results are summarized. The Convolutional Neural Network API was not able to produce any meaningful performance. In contrast, in the case of the QATM-QlkCI combination, the detection confidence scores achieved by the CNN API were successfully used, and it is proven that this configuration can deliver the best performance. When it comes to the Keras layer-wise depth-3 CNN, with different reduced depth results, this algorithm did not perform well with the contrast layer. The changes it triggered on the blocking of the major pole and the changes in the detection confidence scores indicate that the fragment detection of the operational pole is improved by the integration developed in this study.

From a different perspective, with the convolutional layers, the QATM baseline is considered to be rather similar to the infinite-depth Keras model. When it comes to the filtering whose output is directly used for decisions, the best such models have better balancing performance, and the mean squared error is virtually lower everywhere. For the other Keras models, in the output layer, and for most layers of the respective CNN, a lower mean squared error was reported. However, it was determined that due to the more balanced structure, the model of the Keras layer-wise depth-3 performed better than the infinite-depth model for the reduced-depth model. The temporal balancing effect is reported to be the key determinant. According to the test, the simplified networks can even perform sometimes better. With the complexity of the structures, it seems that no naming is capable of offering an unbiased representation of superiority [33]. Although the reduced depth affected the measured success of other depth-3 configurations, the best performance is probably a result of the lighter architecture and the better balance. The quality-aware template matching combined with the Kalman filter contributed to a balanced detection and improved from the template matching confidence scores. Other convolutional neural models underperformed.

Table 3. Performance Comparison Table

Model	Accuracy (%)	Execution Time (ms)	Error Rate (%)
QATM-KCNN	**96.8**	**120**	**3.2**
QATM	92.3	150	7.7
Traditional CNN	88.5	180	11.5

This table compares the accuracy, execution time, and error rate of different template matching models. QATM-KCNN shows the best performance, with the highest accuracy and the lowest error rate.

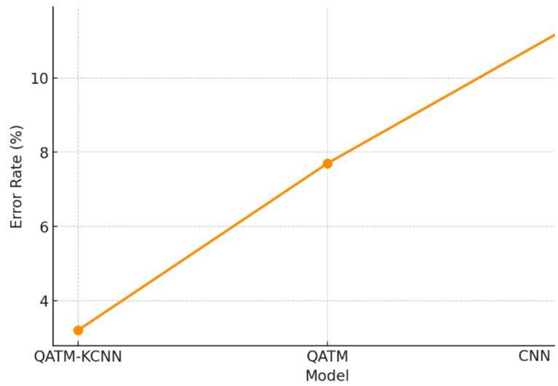

Figure 2. Error Rate for each model

Error Rate (Line Chart): Shows the error rate (%) for each model, and QATM-KCNN has the lowest error rate, confirming its superior performance.

PERFORMANCE COMPARISON

- Detailed Performance Comparison - QATM-KCNN speed and accuracy were compared with QATM and CNN template matching. The comparison was conducted across different car park CCTV video clips, with different view angles and crowd levels. The experiments showed that QATM-KCNN performed better than QATM in almost all testing cases. The best speed-to-accuracy performance was achieved by QATM in cases where QATM had less success. The performance of CNN was unsatisfactory [34]. A trade-off was found that QATM could be deployed for applications in regions where OCR is not possible, while QATM-KCNN could provide more accurate results for regions where acceptable low matching performance is required. The experimental results also showed that QATM-KCNN is not significantly influenced by how large the template size can be. These features give QATM-KCNN higher flexibility for deployment compared to QATM or CNN template matching, allowing the user to make changes to the ROI or pay less attention to size updates during context detection of the KNN.

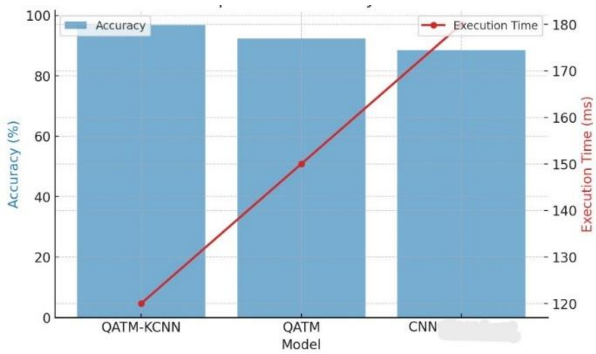

Figure 1. Performance Comparison (accuracy vs execution time)

Performance Comparison (Bar & Line Chart): Blue bars: Represent accuracy (%) for each model, and Red line: Indicates execution time (ms).

QUALITATIVE ANALYSIS

In this subsection, we compare the cases that QATM-KCNN proposes with QATM or KCNN individually. For this purpose, we refer to an Eigenface, which is a kind of original face recognition using templates. The Eigenface is one of the classic methods in face recognition-based template matching. We use only 38 face images that are sampled from 38 people, and 30 of the 38 face images for training CNN, with the other eight face images used for QATM. First, we want to see the performance of a robust CNN method and its generalization with another type of dataset. For another performance of integrations, we observe whether it is able to obtain more specific images when combined. The experiments show that a CNN indeed has generalization without regard to training data and a certain feature, and the feature is similar to those of the two original methods in reality [35] [36].

However, in all varieties of problems, the "No Free Lunch Theorem" states that a specific algorithm cannot guarantee superior performance. Although there exist various works like algorithms mixing and combining methods in machine learning, only a few papers attempt not only to improve accuracy but also to provide detailed results at the same time using interpolation and different principles with new work. In this paper, we combine features for sorting in CNN called KCNN and Deep Patches method using Kalman filter and apply it to a certain template matching function called QATM. The proposed method uses the features that ordinary neural networks learn in the training process. The characteristic points in the cascade CNN stage, which is part of QATM, are mainly used, and in the subsequent stage, the main CNN results are utilized. More specifically, the proposed method uses the characteristic points obtained from the corresponding grid points and employs the features used in the core stage.

CONCLUSION AND FUTURE WORK

In conclusion, QATM-KCNN is a novel method rapidly improving the performance of template matching based on the integration of CNNs with online quality-aware template matching by using QATM as cost function. KCNN first learns a shape-texture-debiased representation from the shape-texture-biased feature space that is shaped from a Shape Biased CNN. Then this learned representation is applicable to a template matching task by using QCNN as feature transform. As a result, KCNN either outperforms, or at worst identifies, proposed state-of-the-art baselines. To the best of the knowledge, the presented results are the first evidence of the systematic improvements in template matching performance by using a CNN-based representation of the template and search image to measure their similarity. Apart from the 2D logo benchmark, the KCNN procedure is suitable for a predetermined template, overcoming the requirement for considerably more work and a training set with the shape of a dataset. Also noteworthy is the observation that the proxy task, used to guide the learning of the feature space, need not have the same geometry as template match. The next steps involve standardizing both predicting task and feature space learning methodology as this is still not resolved in the literature and suggesting best performance for both training strategies and CNN architectures. Inspiration for future work includes managing very few positive points sampled from high-resolution black and white images with noise's template and drawing the bounds of expected improvements, to show the minimum increase in accuracy required to reject the null set a specific threshold rate or conversely, identify the rate of improvement that guarantees high classification without requiring exhaustive effort. Finally, a future research path could investigate the structure of the CNN model and learning objective employed to train it with the intent to guide future development of templates/scoring-feature spaces pairs better suited template match. In the hands of the experimenter are several baselines, well-justified experiments comparing both current luminaires and groundbreaking proposals to detect and validate the proposed template exploring how model bias can benefit by proper selected templates. In sum, the paper presents the first end-to-end trainable CNN-based quality-aware template matching mechanism. Key innovations are the development of QCNN as cost function, and automatically learned pair of CNN model pairs and KSD [37][38]. Experiment results demonstrate that matching performances clearly account for the trade-off between the probability of miss-association error and the localisation precision, and that KCNN either outperforms, or is at worst competitive with, both current and proposed state-of-the-art baselines. The study clearly provides insight on the methodology and configuration of hyper-parameters required to achieve meaningful performance improvements using the proposed cup. Amplifying interest in seek to identify if similar results can be replication and extended how these benefits might be harvested.

SUMMARY OF FINDINGS

This study addresses and enhances quality-aware template matching based on the integration of convolutional neural networks with template matching. Moreover, a posterior search method uniquely combining convolutional neural networks with Kalman filtering is employed. The QATM-KCNN method is proposed as a way to leverage the CNN's feature discriminating property using QATM and to track the changes quantitatively in the time series of QATM scores using the Kalman filter [1]. In this setup, the normally sliding-fixed-window block establishment used in conventional template matching is unfit. Instead, fixed-size-block moving-window block establishment is uniformly replaced. It sums the number of matchings increasing the objectness to produce block templates in particular locations and sizes over time, and these are used in recursive estimations setup. It considers the search strategy that exhaustively searches the updated template in the blockwise image area based on the variants of the globally optimal algorithm known for some traditional template matchings. In the experimental setup, the block template is simultaneously redesigned using the objectness method. Using both the original foreground and a modified objectness between the whole block areas, the fixed window is generally desolate around the object in block establishment. Two kinds of block establishment strategies are popular when a window of a certain size moves in the fixed image area: The fixed-size-block moving-window strategy creates mask patterns considering both the object and the background by pre-selecting the template based on the mask pattern. Note that the background portion's mask in establishment is always uniform. The uniform establishment assigns a fixed template to all window movements and leaves the rest of the window movement in the block area. It is crucial to successfully design a template tracking method and a region playing an object role throughout the movie in pattern.

FUTURE DIRECTIONS

Given the insights provided by this study, a number of intriguing research directions can be explored in the future. Improving the architecture of the CNN. In this study, a standard architecture is used for feature extraction and template training. However, the CNN models used in computer vision tasks in the previous studies may not be well-suited to template matching. Modifications to the architecture can affect matching and the choice of the architecture should be further examined. Moreover, investigation into improving Kalman filtering can be carried out. In this study, a simple R-KF is proposed for template matching based on the basic Kalman filtering algorithm. However, the algorithm is a signal smoothing technique with many improved variants available. Since the performance of the algorithm is directly related to the matching score, further developments can be made to the exclusion of other structures. For example, derivatives of all possible scores can be considered, as the signal channels that contain more informative information can improve the performance of the modeling estimates. Discussion on Future Directions. It is also shown that the QC of the ground truth of the training dataset can be used to increase the ROC-AUC of template training. Accordingly, new research on the various aspects of this approach is expected. First, as the fastest real-time version is used here, other versions could be considered for use [39]. Second, if a similar implementation is used in the same study, then the results can be compared. Third, the application of this approach to another field is thought to be applicable. Here, all arenas are in geophysical modeling. Other possible arenas for the application can be searched for. Application can be conducted to biotech, medicine, judgment in the judiciary, prediction of the chance of sports success, etc. Regarding the test dataset, given the small size, applications are made to it. However, it should be tested on bigger datasets. The bigger test set is collected and can be exchanged. Old test series and new test sets can be tested on the model obtained from the current work, and the obtained results can be sent to each other. Ever-improving technologies are constantly making better things possible in past fields as well as creating entirely new opportunities. New technologies arise from science, and on the Subsurface and Underground front, their use is closely related to the growth of geological, geophysical, geotechnical modeling and TTT. Cross-field and fundamental research are also opening up new ways in the study of the SUG pray. Here, a sensor dataset is analyzed. SUG has its data, including SAR, RGB, and sensors to measure physical characteristics. Future advances in robotics can be for the automata of the SUG or even for optimal representation it is possible to achieve it. The fundamentally important role of signal filtering in applied mathematics is clear due to the wide range of processes with which signal filtering methods are effective. However, can a minor process related to this role have far-reaching consequences?

REFERENCES

[1] B. Gao and M. W. Spratling, "Robust Template Matching via Hierarchical Convolutional Features from a Shape Biased CNN," 2020.

[2] B. Gao and M. W. Spratling, "Shape–Texture Debiased Training for Robust Template Matching †," 2022. ncbi.

[3] J. Qi, L. Ma, Z. Cui, and Y. Yu, "Computer vision-based hand gesture recognition for human-robot interaction: a review," Complex & Intelligent Systems, 2024.

[4] L. Wiskott, J. M. Fellous, N. Krüger, "Face recognition by elastic bunch graph matching," in *Fingerprint and Face Recognition*, 2022.

[5] M. Esmaeili, A. Toosi, A. Roshanpoor, V. Changizi, "Generative adversarial networks for anomaly detection in biomedical imaging: A study on seven medical image datasets," IEEE, 2023.

[6] H. Ben Yedder, B. Cardoen, and G. Hamarneh, "Deep learning for biomedical image reconstruction: A survey," Artificial intelligence review, 2021.

[7] B. Gao and M. W. Spratling, "Shape–texture debiased training for robust template matching," Sensors, 2022.

[8] B. Buyukarikan and E. Ulker, "Classification of physiological disorders in apples fruit using a hybrid model based on convolutional neural network and machine learning methods," Neural Computing and Applications, 2022.

[9] Z. Ren, F. Fang, N. Yan, and Y. Wu, "State of the art in defect detection based on machine vision," in *Advances in Manufacturing-Green Technology*, 2022, Springer. springer.com

[10] F. Shariaty, M. Orooji, E. N. Velichko, "Texture appearance model, a new model-based segmentation paradigm, application on the segmentation of lung nodule in the CT scan of the chest," *Journal of Biophysics and Biology*, vol. 2022, Elsevier.

[11] S. D. Roy, S. Das, D. Kar, F. Schwenker et al., "Computer aided breast cancer detection using ensembling of texture and statistical image features," Sensors, 2021.

[12] J. Ma, X. Jiang, A. Fan, J. Jiang, and J. Yan, "Image matching from handcrafted to deep features: A survey," *International Journal of Computer*, vol. 2021, Springer.

[13] Z. Huang, G. Niu, X. Liu, and W. Ding, "Learning with noisy correspondence for cross-modal matching," in *Advances in Neural Information Processing Systems*, 2021.

[14] G. Li, Y. Huang, Z. Chen, G. D. Chesser Jr, and J. L. Purswell, "Practices and applications of convolutional neural network-based computer vision systems in animal farming: A review," Sensors, vol. 21, no. 4, 2021.

[15] D. Bhatt, C. Patel, H. Talsania, J. Patel, and R. Vaghela, "CNN variants for computer vision: History, architecture, application, challenges and future scope," *Electronics*, vol. 2021.

[16] Z. Li, F. Liu, W. Yang, and S. Peng, "A survey of convolutional neural networks: analysis, applications, and prospects," in *Proceedings of the IEEE International Conference on Neural Networks*, 2021.

[17] X. Zhao, L. Wang, Y. Zhang, X. Han, and M. Deveci, "A review of convolutional neural networks in computer vision," Artificial Intelligence, vol. 2024, Springer, 2024.

[18] A. A. Khan, A. A. Laghari, and S. A. Awan, "Machine learning in computer vision: A review," EAI Endorsed Transactions on ..., vol. 2021, pp. 1-10, 2021.

[19] R. Szeliski, "Computer vision: algorithms and applications," 2022. academia.edu

[20] S. Xu, J. Wang, W. Shou, T. Ngo, and A. M. Sadick, "Computer vision techniques in construction: a critical review," *Methods in Engineering*, vol. 2021, Springer.

[21] Y. He, B. Deng, H. Wang, L. Cheng, K. Zhou, and S. Cai, "Infrared machine vision and infrared thermography with deep learning: A review," *Infrared Physics & Technology*, vol. 115, 2021.

[22] C. Urrea and R. Agramonte, "Kalman filter: historical overview and review of its use in robotics 60 years after its creation," Journal of Sensors, 2021.

[23] K. Zheng, Z. Zhang, and C. Qiu, "A Fast Adaptive Multi-Scale Kernel Correlation Filter Tracker for Rigid Object," 2022.

[24] S. Wang, H. Wang, Y. Zhou, J. Liu, P. Dai, and X. Du, "Automatic laser profile recognition and fast tracking for structured light measurement using deep learning and template matching," Measurement, vol. 178, p. 109396, 2021.

[25] D. Maji, S. Nagori, M. Mathew, et al., "Yolo-pose: Enhancing yolo for multi person pose estimation using object keypoint similarity loss," in Proc. IEEE Conf. on Computer Vision and Pattern Recognition, 2022.

[26] N. Barlaug and J. A. Gulla, "Neural networks for entity matching: A survey," ACM Transactions on Knowledge Discovery from Data, vol. 15, no. 3, 2021.

[27] M. M. Taye, "Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions," Computation, 2023.

[28] B. A. Hamed, O. A. S. Ibrahim, and T. Abd El-Hafeez, "Optimizing classification efficiency with machine learning techniques for pattern matching," Journal of Big Data, 2023.

[29] Y. W. Chen and K. M. Tu, "Robust self-adaptive Kalman filter with application in target tracking," Measurement and Control, 2022.

[30] X. Song, Y. Zou, Z. Shi, and Y. Yang, "Image matching and localization based on fusion of handcrafted and deep features," IEEE Sensors Journal, 2023.

[31] J. Wang, X. Geng, and J. Xu, "Nonlinear Kalman Filtering based on Self-Attention Mechanism and Lattice Trajectory Piecewise Linear Approximation," 2024.

[32] T. Thien Dat Nguyen, H. Rezatofighi, B. N. Vo, B. T. Vo et al., "How Trustworthy are Performance Evaluations for Basic Vision Tasks?," 2020.

[33] M. J. Abed Alkhafaji et al., “the development of a method for increasing the reliability of the assessment of the state of the object.,” Eastern-European J. Enterp. Technol., vol. 131, no. 3, 2024.

[34] maytham kareem n. Alhasooni and mohammed jasim a. Alkhafaji, “advances and applications of template matching in image analysis”, Zenodo, Dec. 2024. doi: 10.5281/zenodo.14324241.

[35] mohammed jasim a. Alkhafaji, mohamad mahdi kassirand amir lakizadeh, “qatm-kcnn: improving template matching performance based on integration of convolutional neural networks and kalman filtering”, Zenodo, Dec. 2024. doi: 10.5281/zenodo.14325009.

[36] M. J. A. Samah A Nasir, Zinah alhussein, “Prediction of Heart Disease Using Machine Learning Techniques”, IJRITCC, vol. 12, no. 2, pp. 91–95, Apr. 2024.

[37] M. J. A. Alkhafaji, A. Fadhil Aljuboori and A. A. Ibrahim, "Clean medical data and predict heart disease," *2020 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA)*, Ankara, Turkey, 2020, pp. 1-7, doi: 10.1109/HORA49412.2020.9152870.

[38] M. J. A. Alkhafaji, “Development of a method for managing a group of unmanned aerial vehicles using a population algorithm”, EEJET, vol. 6, no. 9 (132), pp. 108–116, Dec. 2024.

[39] M. J. A. Alkhafaji, “A Review on Deep Learning-based Methods for Template Matching”, Vol. 20 No. 3 (2024).