*LatIA*. 2025; 3:120
doi: 10.62486/latia2025120
ISSN: 3046-403X

ORIGINAL

# Robust Face Tracking Under Challenging Conditions Using Linear Regression and YOLO algorithm

## Seguimiento facial robusto en condiciones desafiantes mediante regresión lineal y algoritmo YOLO

Haya alhadramy[1], Mazen Alzyoud[1], Mohammad Subhi Al-Batah[2], Najah AL-shanableh[1]

[1]Al al-Bayt University, Department of Computer Science, Faculty of Prince Al-Hussein Bin Abdallah II for IT, 2513 Mafraq, Jordan.
[2]Jadara University, Department of Computer Science, Faculty of Science and Information Technology, Irbid, Jordan.

**Cite as**: Alhadramy H, Mazen A, Al-Batah MS, Al-shanableh N. Robust Face Tracking Under Challenging Conditions Using Linear Regression and YOLO algorithm. LatIA; 3:120. https://doi.org/10.62486/latia2025120

**Submitted:** 19-03-2024 **Revised:** 08-06-2024 **Accepted:** 08-10-2024 **Published:** 01-01-2025

**Editor:** Dr. Rubén González Vallejo

**Corresponding author:** Mohammad Subhi Al-Batah

**ABSTRACT**

Face detection and tracking play a crucial role in various computer vision applications, including surveillance, fault face detection systems, artificial intelligence, etc. The objective of this paper is to enhance the precision of face detection and tracking through the introduction of an innovative approach centered on the linear regression algorithm. The effectiveness of the proposed method was compared to the traditional Kalman filter approach. Additionally, the study explored the integration of the YOLO algorithm for face detection with the linear regression tracking algorithm to further enhance accuracy. The proposed algorithm's performance is assessed through comprehensive experiments on annotated images and video sequences affected by occlusions or other issues such as poor lighting conditions and motion blur. These experiments utilize the COCO dataset, operating at a speed of 60 FPS. The experimental results show that the proposed method can accurately track the human face in different facial positions.

**Keywords:** Face Detection; Face Tracking; YOLO Algorithm; Linear Regression; Kalman Filter.

**RESUMEN**

La detección y el seguimiento de rostros desempeñan un papel crucial en diversas aplicaciones de visión por computadora, incluida la vigilancia, los sistemas de detección de rostros defectuosos, la inteligencia artificial, etc. El objetivo de este documento es mejorar la precisión de la detección y el seguimiento de rostros mediante la introducción de un enfoque innovador centrado en sobre el algoritmo de regresión lineal. La eficacia del método propuesto se comparó con el enfoque tradicional del filtro de Kalman. Además, el estudio exploró la integración del algoritmo YOLO para la detección de rostros con el algoritmo de seguimiento de regresión lineal para mejorar aún más la precisión. El rendimiento del algoritmo propuesto se evalúa mediante experimentos exhaustivos con imágenes anotadas y secuencias de vídeo afectadas por oclusiones u otros problemas como malas condiciones de iluminación y desenfoque de movimiento. Estos experimentos utilizan el conjunto de datos COCO, operando a una velocidad de 60 FPS. Los resultados experimentales muestran que el método propuesto puede rastrear con precisión el rostro humano en diferentes posiciones faciales.

**Palabras clave:** Detección de Rostros; Seguimiento de Rostros; Algoritmo YOLO; Regresión Lineal; Filtro de Kalman.

## INTRODUCTION

Recently, Face tracking has become a field that has a lot of research and development. Face tracking is applied to track the movement of target objects, such as human tracking, moving target tracking, security, and traffic surveillance. Particularly, human tracking is often used in real life with many applications such as robot vision, monitoring, and so on. Any human tracking algorithm needs to solve detection and tracking problems across multiple frames.[1] However, the drawbacks of algorithms are the ability to work in complex backgrounds, light sources, fails to detect any faces, indicating a deficiency in its ability to identify faces accurately and many other different conditions is also a challenge that attracts many researchers.

Recently, a lot of deep-learning approaches have been developed to boost accuracy. The merit of this approach is that it automates feature extraction and has high flexibility in detecting objects with high accuracy. This approach needs data to train and improve accuracy.[1] The You Only Look Once (YOLO) algorithm has recently been the best object detection model; its algorithm has high speed and accuracy. The YOLO algorithm is simple, speedy and refreshing. It predicts what objects are present and where they are.[2] The R-CNN[3], Fast R-CNN[4], and Faster R-CNN are good detection models. However, due to a two-stage network, it could be faster.[5] Rudolf E. Kalman proposed the linear quadratic estimate (LQE) method in 1960. It is a tracking algorithm to improve accuracy compared to just one measurement result, and the Kalman filter uses a sequence of measurement values that are affected by noise or error to predict the target position. The next frame's position and velocity estimation errors are reduced via the Kalman filter.[1]

The linear regression is a statistical test applied to a data set to identify and quantify relationships among the variables under consideration.[6] The univariate statistical tests such as chi-square, t-test, and analysis of variance (ANOVA) regardless of the effect of other variables or confounders in the analysis.

Linear regression has two types: Simple linear regression and multiple linear regression simple linear regression has one independent variable x, and one dependent variable y, where $y$ = ß0 +ß1, the influence of independent variables and the interaction of dependent variables are separated by simple regression.[7]

### Literature review

Machine learning is the capacity of a system to automatically gather, integrate, and then develop knowledge from enormous amounts of data and then expand the learned information on its own without being explicitly programmed to do.[8]

Neural networks are a great tool for both image classification and quick detection of the desired items. Deep learning or CNNs are being used far more often, and they have a wide range of uses, including tracking systems, facial recognition, and people detection.[9] Above all, people-tracking systems are essentially able to identify and monitor the movement of people within a camera frame.[10] Previous studies indicated that neural networks would perform better if they had access to a large, diversified database for self-training for instance, the system would function more effectively if there were photographs in the training database featuring multiple people.[11]

In 1894 Sir Francis Galton proposed an algorithm called Linear Regression for statistical test applied to a data set to identify and quantify relationships among the variables under consideration.[6]

The proposed Linear Regression and Adaptive Appearance Models for fast simultaneous modelling and tracking is studied in [12]. The study focuses on estimate a face's position and motion parameters; the simplicity and interpretability of linear regression make it an attractive alternative to more complex methods, its algorithm leverages training data to learn the relationship between facial features and the target variables, enabling accurate estimation of face positions.

The YOLO technique is distinct from previous object detection algorithms in that it makes predictions about the bounding boxes and their class probabilities using a single neural network.[2] Its performance can therefore be immediately optimized end-to-end, and the input can be computed very quickly afterward. An important part of face detection metabolism has been played by linear regression face detection technology becomes more useful with increased speed, accuracy, and robustness.[13]

Various algorithms were proposed for face detection and tracking. The face moving person completely disappears from the algorithm's view, which could occur due to occlusions, sudden movements, or changes in lighting conditions. After a hidden face reappears, the current algorithm often considers it a new face and assigns a new identifier, leading to inconsistency in tracking and loss of identity. This inconsistency reduces the system's accuracy and challenges maintaining a reliable and continuous tracking process Therefore, this paper proposes a combination of face detection based on the YOLO algorithm and linear regression for face tracking.

## METHOD

In this study, the face detection and tracking process consists of several steps. The overall procedure can be summarized as follows: Fig. 1 depicts an overview of the multiple face-tracking systems used in this study. As shown in the diagram, the video begins with a sequence of images import frames from coco dataset to

https://doi.org/10.62486/latia2025120 ISSN: 3046-403X

determine whether Yolo detects the face or not. If Yolo cannot detect the face, use linear regression to predict the face's future location and then update the face location; however, if Yolo detects the face, use linear regression to see if the face already exists and is saved in the list; otherwise, create a new linear tracker and save it in the list, then predict the face's following location and finally update the face location.

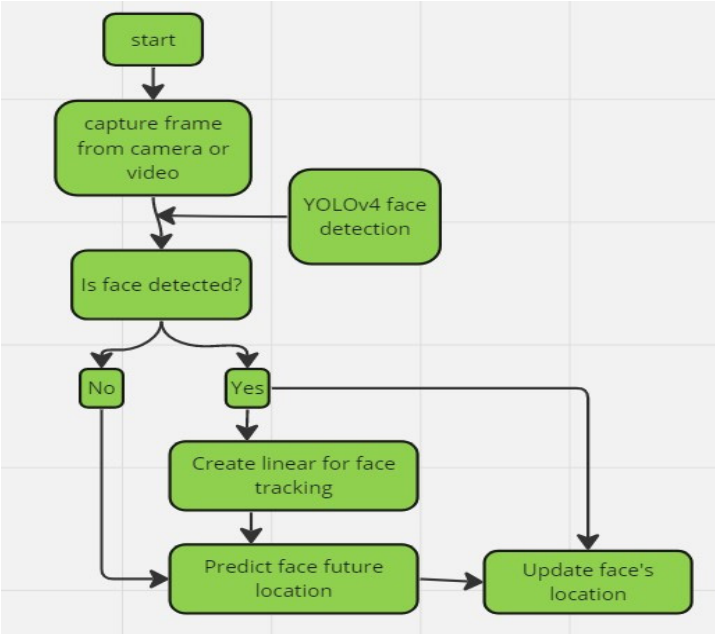


**Figure 1.** Flow chart of the proposed algorithm

A. Face detection based on YOLO algorithm

represents a single convolutional network that simultaneously predicts multiple bounding boxes and class probabilities for those boxes, YOLO trains on full frames and directly optimizes detection performance, this unified model has several advantages over traditional object detection methods.

B. Face tracking based on linear regression.

Simple linear regression is a statistical term for predicting relationships between a single dependent variable Y, and one independent variable. There is only one independent variable X, but more than one independent variable is called multiple linear regression.(14)

$Y = \beta 0 + \beta 1X$ …… (1)

At the beginning of the first frame, the linear regression is detected after three initial frames to prevent false alarms. Once detected, the linear regression takes the box in the accepted format (Xmin, Ymin, Xmax, Ymax, width, height). This box is then saved in the history for each track, and the regression returns both the original detected box and a prediction with all variables set to zero. This is because it is the first detection for each object, and there has yet to be past data.

After the initial detection, the detection is partitioned into four lists, and four threads are generated for projecting subsequent locations for each variable through linear regression. Following this, all predictions from linear regression are amalgamated into a unified list, which is stored in a static list and subsequently utilized to generate a bounding box.

In subsequent frames, the detection for each object arrives in the format (MinX, MinY, MaxX, MaxY). Following this, the detection is divided into four lists. Utilizing a static list containing prior detections, the new variable value is combined with the previous one. Four threads are generated using linear regression to predict the subsequent location for each variable, followed by aggregating the predictions from linear regression into a unified list. This list is appended to the static list and contains up to five previous detections. When the static list reaches a count of five, the earliest prediction is removed, the new one is added, and the resulting single list is returned to draw a bounding box.

Subsequently, the point list is fed into the linear regression model for individual variables, facilitating the training of this model. The linear regression model can forecast the forthcoming location for each frame.

## RESULTS

The experiment is implemented by using Python -V (3.10.11- 64 bit) via Visual Studio Code V 1.73.1 installed on the device has an Intel(R) Core (TM) i5-1035G1 CPU @ 1.00GHz 1.20 GHz, 8 GB RAM.

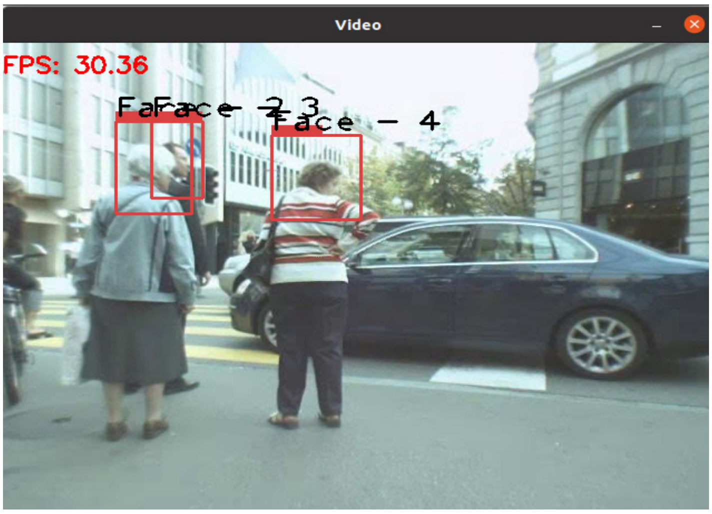


**Figure 2.** Result of proposed method

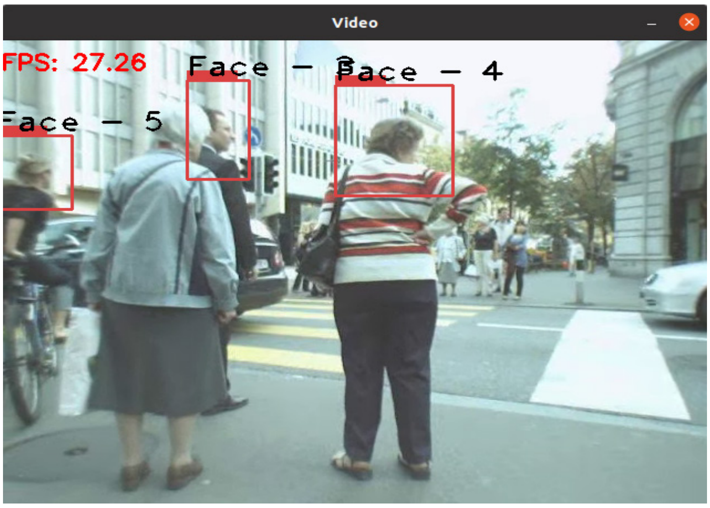


**Figure 3.** Result of proposed method

Figure 2 and figure 3 represent the location of faces detected by the YOLOv4 algorithm. The predicted face location by linear regression to clearly visualize the proposed model, the study overlaid all the above results in multiple images and showed the trajectory of the face. When the faces are detected, linear regression initially predicts its state in the present images frame from the camera.


**Figure 4.** Result of proposed method

https://doi.org/10.62486/latia2025120 ISSN: 3046-403X

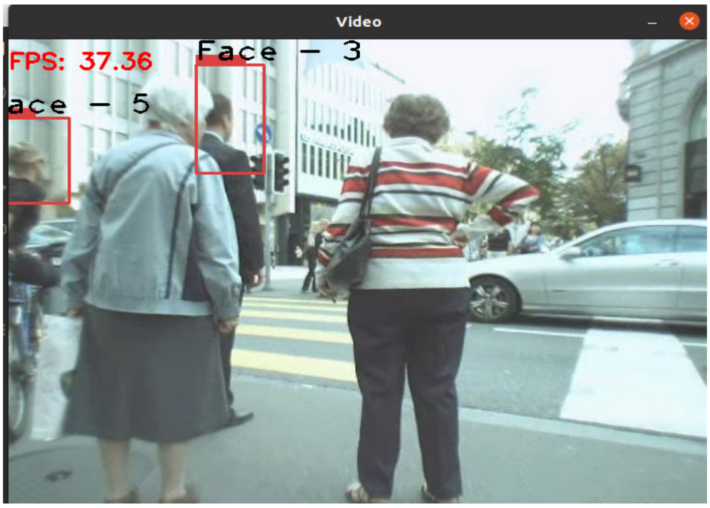


**Figure 5.** Result of proposed method

In figure 5 the complete disappearance of a face was discovered and in figure 6 When the face is missing, linear regression only depends on its previous state to predict the current position of the face and keep the ID. Table 1 shows the experiment comparison for face detection and tracking accuracy

**Table 1.** Face Detection and Tracking Accuracy

| Metric | Kalman filter with YOLO | Linear regression with YOLO |
|---|---|---|
| Accuracy | 85 % | 90 % |
| Precision | 83 % | 86,8 % |
| Recall | 81,4 % | 84,7 % |
| F1 Score | 82,2 % | 85,5 % |

The result shows that the face detection and tracking accuracy in multiple faces tracking in face occluded and other conditions is reached to 90 %.

## CONCLUSION

The present study aimed to improve the accuracy of multiple face detection and tracking using a linear regression algorithm compared to the traditional Kalman filter approach. The effectiveness of the proposed model was evaluated based on several key objectives. The primary objective was to assess and compare face detection and tracking accuracy and performance using the linear regression algorithm and the Kalman filter and the impact of various factors, such as lighting conditions in good light, on the performance of the linear regression algorithm was analyzed, the YOLO algorithm was integrated with the linear regression algorithm to assess its impact on face detection accuracy in real-time live video.

## REFERENCES

1. Nguyen, V. T., Chu, D. T., Phan, D. H., & Tran, N. T. (2023). An Improvement of the Camshift Human Tracking Algorithm Based on Deep Learning and the Kalman Filter. Journal of Robotics, 2023.

2. Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).

3. Bharati, P., & Pramanik, A. (2020). Deep learning techniques—R-CNN to mask R-CNN: a survey. Computational Intelligence in Pattern Recognition: Proceedings of CIPR 2019, 657-668.

4. Wan, S., & Goudos, S. (2020). Faster R-CNN for multi-class fruit detection using a robotic vision system. Computer Networks, 168, 107036.

5. Cao, C., Wang, B., Zhang, W., Zeng, X., Yan, X., Feng, Z., ... & Wu, Z. (2019). An improved faster R-CNN for small object detection. Ieee Access, 7, 106838-106846.

6. Kumari, K., & Yadav, S. (2018). Linear regression analysis study. Journal of the practice of Cardiovascular Sciences, 4(1), 33.

7. Maulud, D., & Abdulazeez, A. M. (2020). A review on linear regression comprehensive in machine learning. Journal of Applied Science and Technology Trends, 1(4), 140-147.

8. Sen J, Mehtab S. (2021) A comparative study of optimum risk portfolio and eigen portfolio on the Indian stock market. International Journal of Business Forecasting and Marketing Intelligence. Inderscience, Paper ID: IJBFMI-90288.

9. N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," Proc. - Int. Conf. Image Process. ICIP, vol. 2017-Septe, pp. 3645–3649, 2018, doi: 10.1109/ICIP.2017.8296962.

10. R. Muñoz-Salinas, R. Medina-Carnicer, F. J. Madrid-Cuevas, and A. Carmona-Poyato, "Multi-camera people tracking using evidential filters," Int. J. Approx. Reason., vol. 50, no. 5, pp. 732–749, 2009, doi: 10.1016/j.ijar.2009.02.001.

11. Azhar, M. I. H., Zaman, F. H. K., Tahir, N. M., & Hashim, H. (2020, August). People tracking system using DeepSORT. In 2020 10th IEEE international conference on control system, computing and engineering (ICCSCE) (pp. 137-141).

12. Ellis, L., Dowson, N., Matas, J., & Bowden, R. (2011). Linear Regression and Adaptive Appearance Models for Fast Simultaneous Modelling and Tracking. International Journal of Computer Vision, 95(23), 154- 179.

13. Gupta, I., Mittal, H., Rikhari, D., & Singh, A. K. (2022). Mlrm: A multiple linear regression-based model for average temperature prediction of a day. arXiv preprint arXiv:2203.05835.

14. Aung, H., Valentinovich, B. A., & Aye, B. (2022, May). Real-Time Face Tracking Based on the Kalman Filter. IEEE In 2022 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM) (pp. 842-846).

15. Sadura, P. (2021, September). Motion-based multiple object detection and tracking in video. IEEE In 2021 Signal Processing Symposium (SPSympo) (pp. 248-251).

16. M. S. Al-Batah, M. Alzyoud, R. Alazaidah, M. Toubat, H. Alzoubi, & A. Olaiyat, "Early Prediction of Cervical Cancer Using Machine Learning Techniques," Jordanian Journal of Computers and Information Technology, 8(4), 2022.

17. M. S. Al-Batah, "Ranked features selection with MSBRG algorithm and rules classifiers for cervical cancer," Int. J. online Biomed. Eng., vol. 15, no. 12, pp. 4–17, 2019, doi: 10.3991/ijoe.v15i12.10803.

18. M. Al-Batah, B. Zaqaibeh, S. A. Alomari, and M. S. Alzboon, "Gene Microarray Cancer classification using correlation based feature selection algorithm and rules classifiers," Int. J. online Biomed. Eng., vol. 15, no. 8, pp. 62–73, 2019, doi: 10.3991/ijoe.v15i08.10617.

19. A. Quteishat, M. Al-Batah, A. Al-Mofleh, and S. H. Alnabelsi, "Cervical cancer diagnostic system using adaptive fuzzy moving k-means algorithm and fuzzy min-max neural network," J. Theor. Appl. Inf. Technol., vol. 57, no. 1, pp. 48–53, 2013.

20. M. S. Al-Batah, A. Zabian, and M. Abdel-Wahed, "Suitable features selection for the HMLP network using circle segments method," Eur. J. Sci. Res., vol. 67, no. 1, pp. 52–65, 2011.

21. M. S. Al-Batah, "Automatic diagnosis system for heart disorder using ESG peak recognition with ranked features selection," International Journal of Circuits, Systems and Signal Processing, 13(June), 391–398, 2019.

22. M. S. Al-Batah, M. S. Alkhasawneh, L. T. Tay, U. K. Ngah, Hj Lateh, H., and N. A. Mat Isa, "Landslide Occurrence Prediction Using Trainable Cascade Forward Network and Multilayer Perceptron. Mathematical Problems in Engineering," 2015, https://doi.org/10.1155/2015/512158

https://doi.org/10.62486/latia2025120 ISSN: 3046-403X

23. M. S. Alkhasawneh, U. K. Ngah, L. T. Tay, N. A. Mat Isa, and M. S. Al-Batah, "Determination of important topographic factors for landslide mapping analysis using MLP network," The Scientific World Journal, 2013, https://doi.org/10.1155/2013/415023

24. A. F. karimBaareh, A. Sheta, and M.S. Al-Batah, "Feature based 3D Object Recognition using Artificial Neural Networks," International Journal of Computer Applications, 44(5), 1–7, 2012, https://doi.org/10.5120/6256-8402

## FINANCING
This work is supported from Jadara University under grant number [Jadara-SR-Full2023], and Al al-Bayt University, Mafraq, 25113, Jordan.

## CONFLICT OF INTEREST
The authors declare that there is no conflict of interest regarding the publication of this paper.

## AUTHORSHIP CONTRIBUTION
*Conceptualization:* Mohammad Subhi Al-Batah, Haya alhadramy, Najah AL-shanableh, Mazen Alzyoud.
*Data curation:* Haya alhadramy, Najah AL-shanableh, Mazen Alzyoud.
*Formal analysis:* Haya alhadramy, Najah AL-shanableh, Mazen Alzyoud.
*Research*: Haya alhadramy, Najah AL-shanableh, Mazen Alzyoud.
*Methodology:* Haya alhadramy, Najah AL-shanableh.
*Project management:* Najah AL-shanableh, Mazen Alzyoud.
*Resources:* Haya alhadramy, Najah AL-shanableh.
*Software*: Haya alhadramy.
*Supervision:* Najah AL-shanableh, Mazen Alzyoud.
*Validation:* Mohammad Subhi Al-Batah, Haya alhadramy, Najah AL-shanableh, Mazen Alzyoud.
*Display:* Haya alhadramy.
*Drafting - original draft:* Haya alhadramy.
*Writing - proofreading and editing:* Mohammad Subhi Al-Batah, Haya alhadramy, Najah AL-shanableh, Mazen Alzyoud.