DOI: 10.15849/IJASCA.230320.11

Int. J. Advance Soft Compu. Appl, Vol. 15, No. 1, March 2023
Print ISSN: 2710-1274, Online ISSN: 2074-8523

Multi-Object Tracking Algorithm for Poultry Behavior Anomaly Detection

Jasmine Khairunissa[1], Sri Wahjuni[1*], Iman Rahayu H. Soesanto[2], Wulandari[1], Auriza R. Akbar[1], Hendra Rahmawan[1]

[1] Department of Computer Science Faculty of Mathematics and Natural Science, IPB University
*email: my_juni04@apps.ipb.ac.id
2 Department of Animal Production and Technology, Faculty of Animal Science, IPB University

Abstract

One of the five freedom principles in animal welfare is freedom from pain, injury, and diseases. Chickens that don't move for a long time are more likely to feel uncomfortable with the environment or with their own body. Observing each poultry manually will waste a lot of time and effort of the farmer, therefore it is necessary to have a way to monitor the poultry. One way that can be done is to develop a chicken movement monitoring system using the Multi-Object Tracking (MOT) algorithm. The MOT algorithm is an object tracking method consisting of object detection and tracking. In previous studies, a hybrid method was used which only detects objects every few frames and tracks objects in that period alternately with the object detection stage. This method produces a Multi-Object Tracking Precision (MOTP) score of 60.4%. Meanwhile, in this study a sequential method is used, where the program performs object detection and object tracking stage sequentially, which results in a MOTP score of 87.64%. Moreover, to detect chickens that have not moved for a long time, the Isolation Forest anomaly algorithm is used. The results of this study can integrate into a real-time chicken coop remote visual monitoring.

Keywords: *animal welfare, broiler, object detection, i-Forest, real-time monitoring.*

1 Introduction

The consumption of purebred and native chicken meat in Indonesia continues to increase, which causes poultry farming in Indonesia to be increasingly in demand by the public in Indonesia. This is shown by the presence of 401 poultry companies that are active in production and meet the requirements with a total of 18,835

Received 4 November 2022; Accepted 8 February 2023

workers in 2019 [1]. In the same year, these livestock companies earned a total income of IDR 13.46 trillion. The total expenditure that must be incurred by breeders is IDR 7.27 trillion, which includes 66.2% of the cost of feed; 10.96% of workers' wages; 6.06% of day old chicks (DOC) purchases; 4.40% electricity and water; 4.19% drugs; 1.10% fuel oil; and 7.07% other costs [1]. Broiler chickens have entered the industry and have been produced in large quantities such as nuggets, sausages, and chicken meatballs [2]. The large amount of income and the popularity of chicken farming in Indonesia do not mean that there are no obstacles. A couple of problems that still need to be addressed are business efficiency and livestock welfare.

The production efficiency of broiler farm was analyzed in a study conducted by [3] in Kampar Regency, which is the center of broiler chicken production in Riau Province. The research shows that the number of livestock companies that were not yet efficient was more than the number of livestock companies that have been doing their business efficiently, which meant that the majority of farmers have not been able to optimize inputs to produce maximum output and profit.

It can be seen from the breakdown of expenditure costs detailed by [1] that feed costs account for more than half of the total expenses of poultry companies. Meanwhile, research by [4] examined the application of animal welfare for broilers in Bangka Regency by modifying one of the indicators of the animal welfare assessment method, the Animal Needs Index (ANI-35). In this study, indicators of locomotion, social interaction, light, noise, quality of shelter, and quality of human care for animals were used, with an ANI index score of 16, which indicates that the welfare of livestock is quite adequate. However, if spatial conditions and other conditions that can affect stress levels, both physically and mentally, are not met for a long period of time, then this index score cannot be considered valid [5][6]. Another animal welfare assessment method is the Welfare Quality® (WQ) protocol which includes the following principles: good feeding, good housing, good health, and appropriate behavior [7].

The principle of freedom in animal welfare known as the 5F consists of: freedom from hunger and thirst; freedom from discomfort; freedom from pain, injury or diseases; freedom from fear and distress; and freedom to express natural behavior [8]. Genetic selection and prolonged molecular engineering that occurs in broiler chickens cause rapid muscle growth but not matched by organ growth causing broilers to be susceptible to health problems, such as metabolic disorders Sudden Death Syndrome (SDS), abnormalities in the legs, and decreased activity that causes weakness limbs [9]. The lameness and weakness of the legs make the birds spend long periods of time in a sitting position. If the quality of the mat used is poor, it can cause bruising on the poultry skin such as hock burns, foot pad dermatitis (FPD), and blisters in the poultry chest area [10]. From these problems, it can be concluded that monitoring and detecting the anomalies in the movement behavior in livestock is very important to develop.

Through image and video technology, livestock monitoring can be performed anywhere. In analyzing image and video data, additional steps are needed in order to use the information. These stages can be in the form of feature extraction, image classification, object detection, and object tracking. Detecting anomalies in livestock movement behavior can be done with the help of the multi-object tracking algorithm, Multi-Object Tracking (MOT). The MOT algorithm itself is one of the challenges in the field of computer vision which aims to estimate the direction of movement of the observed object [11]. MOT has been implemented for tracing construction workers [12], road users, and vehicles [13]. This algorithm works by combining object detection methods with object identity associations. The object detection model used will affect the results of the object tracking model [14]. Object detection algorithms are broadly divided into 2 groups: regional proposal-based approaches and regression/classification approaches. The review conducted by Zhao et al. [15] against object detection algorithms shows that the Single Shot Multi-Box Detector (SSD) algorithm by Liu et al. [16] has better performance when compared to other object detection algorithms. SSD algorithms have been implemented to detect various objects such as vehicles [17], communication network insulators in aerial imagery [18], building damage [19], road users [20], hand gestures [21], the condition of dairy cattle [22], and the use of face masks [23].

Previous research works [24] implemented the MOT algorithm and Single Shot Multi-Box Detector (SSD) for detecting the movement of chickens, which resulted in object movement plots. However, the resulting precision value was still low. This is most likely caused by the exchange of identities of the intersecting objects. To overcome this problem in this study, the Simple Online and Real-time Tracking (SORT) algorithm is used for object identity associations and the labeling of the identity of each object in the data at the evaluation stage. In addition, this study also analyzes the detection time of each chicken object to immediately identify when an anomaly occurs using the isolation Forest (iForest) algorithm [25]. The anomaly observed in this study is an object with a longer detectable time when compared to the surrounding objects, because objects are more easily detected at rest. The iForest algorithm was chosen because of its ability to quickly recognize data anomalies and use less space than the Support Vector Machine (SVM), Local Outlier Factor (LOF), and K-Means algorithms [26].

2 Related Works

2.1. Animal Welfare

Animal welfare is a term that refers to the state of the individual in relation to its environment [27]. Animals, like other living things, have needs that must be met in order to survive. These needs influence the response and behavior of the animal. This definition of welfare has several implications: welfare is a characteristic of an animal, not something that can be given away; the level of well-being varies from “very bad” to “very good”; well-being can be measured scientifically; measurement of the animal's failures and difficulties in survival can provide information about how bad the level of welfare is; general knowledge of the condition of the animal must be accompanied by an assessment of the condition of the animal to study and improve its welfare; animals have different methods of survival that must be viewed from various sides, for example, one of the assessment indicators used states that a good level of welfare is not necessarily a good level of overall well-being when viewed from the other side [28]. The Five Freedoms (5F) is the principle of animal welfare under human care, which consists of: freedom from hunger and thirst, freedom from discomfort, freedom from pain, injury or disease, freedom from fear and distress, and freedom to express natural behavior [8]. This principle was developed by the UK Farm Animal Welfare Council and has been adopted by the World Organization for Animal Health (OIE). The methods that can be used to measure the level of animal welfare are the Animal Needs Index (ANI) [5] and the Welfare Quality® (WQ) protocol [7].

2.2. Single Shot Multi-Box Detector (SSD)

Single Shot Multi-Box Detector (SSD) is an object detection method that uses a single deep neural network [16]. SSDs are built based on a feed-forward convolutional network that generates a set of bounding boxes and scores for each detected object. The SSD architecture consists of a basic network, an additional network in the form of a convolution layer, and a Non-Maximum Suppression (NMS) stage. At the beginning of the formation of SSD, the basic network used was the VGG network [29]. However, this basic network can be replaced with other convolution-based networks such as AlexNet [30], GoogLeNet [31], and ResNet [32]. After the basic network, there is an additional network in the form of a convolution layer to detect objects of various sizes. Each feature map in the convolution layer with size $m * n$ can produce as many as $m * n * b$ detections, where b is the number of default boxes that have been set for each feature map. With the number of feature maps = M, the default box scale calculation can be seen in (1) [16].

$$s_k = s_{min} \frac{s_{max} s_{min}}{M-1} (k-1),\ k \in [1, M] \tag{1}$$

In [16] and [18] the recommended number of feature maps is 6 with a convolution kernel size of 3 x 3. The detection results are then processed using NMS to simplify the bounding box detection results.

Parameters that must be considered in SSD implementation are the limits of the Intersection over Union (IoU), confidence, learning rate, momentum, and batch size. IoU is the comparison value between the overlap area of the actual object (ground-truth box) and the default box with the combined area. The IoU calculation can be seen in (2) [33].

$$IoU = \frac{DefaultBox \cap GroundTruthBox}{DefaultBox \cup GroundTruthBox} \quad (2)$$

2.3. Multi-Object Tracking (MOT)

Multi-object Tracking (MOT) is the process of localizing many objects that move within a certain period of time from data taken using a camera or other recording device [14]. The increasing number of available image and video data and the increasing need for object tracking have made the MOT algorithm develop in recent years, including the use of deep learning [34] and the development of the FairMOT algorithm [11].

The algorithms developed for solving MOT problems are very diverse and can be distinguished based on the initialization method, processing mode, and output type [35]. Based on the initialization method, MOT is divided into 2 categories: Detection-based Tracking (DBT) and Detection-free Tracking (DFT). DBT is widely used on specific target objects such as road users [36], vehicles [15], to goats [37]. While DFT detects all objects that appear (not specific). In terms of processing mode, MOT is divided into online and offline browsing. Processing online searches is done sequentially while in offline searches processing is done using batches of data frames used. The results issued by the MOT are also divided into two: deterministic – when the results are fixed even though the process is repeated many times – and probabilistic. This is influenced by the use of deterministic optimization and probability inference.

The evaluation used is the Multi-Object Tracking Precision (MOTP) value obtained by comparing the error value with the total number of matches and misses in the object tracking process [38]. Objects are categorized as match when the bounding box coordinates of the search results are the same as the reference data bounding box coordinates and are categorized as miss if they do not meet these requirements. The MOTP equation can be seen in (3).

$$MOTP = \frac{\sum_{i,t} d_t^i}{\sum_t c_t} \quad (3)$$

2.4. Simple Online and Realtime Tracking (SORT)

The Simple Online and Realtime Tracking (SORT) method is an object tracking method [39] for the Convolutional Neural Network (CNN)-based object detection method using the Kalman filter and the Hungarian Algorithm. This algorithm works by forming an estimation model of the position of each object as x with u, v as the coordinates of the center point (centroid) of the object and s, r as the scale and ratio of the bounding box of each object as in (4).

$$x = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T \quad (4)$$

The object then goes through the IOU distance assessment between the detected object and all the predicted bounding boxes generated. An object is considered lost when it is not detected during a T_{lost} frame. In this study, the T_{lost} = 15 value was used according to the video data frame rate. This algorithm works at a high speed when compared to other search algorithms [39]. The main components of this method are detection, spreading the position of the object to the next frame, association of the currently detected object with the previous object, and maintaining the detected object by establishing tracks that will be deleted if the object is not detected during the T_{lost} frame. It is different from the centroid tracking algorithm which only uses Euclidean Distance to determine the distance between objects between frames [40].

2.5. Isolation Forest (I-Forest)

The Isolation Forest (iForest) method is an anomaly detection method that focuses on isolating anomalies [25]. This method works by generating a collection of trees (iTree) from the data and separating the anomalies that have the shortest path length. The variables used are the number of trees created and the size of the sampling. The number of trees and the recommended sampling size are 100 and 256, respectively. iForest consists of two stages: training and testing. In the training phase, iTree is generated by dividing the training data recursively until the data is isolated or reaches the maximum height. Then in the testing phase, the anomaly score is calculated and used to sort the anomaly. Where n is the number of data, $h(x)$ is the path length, $c(n)$ is the average $h(x)$ of all data, and *E(h(x))* is the average $h(x)$ of a set of isolated trees, anomaly score can be calculated using (5) [25]. When a data has an anomaly score close to -1, then the data is definitely an anomaly and when the anomaly score close to 1, then the data can be classified as normal data.

$$s(x, n) = 2^{-\frac{E(h(x))}{c(n)}} \quad (5)$$

2.6. MOT Implementation

In a previous study [24], the MOT algorithm was applied to detect the movement of poultry objects using SSD as an object detection method and centroid tracking algorithm as an object tracking method between frames. This study uses the framework for the application of the MOT algorithm for the detection of human movement [42] with a combination method of object detection and object tracking. With this method, object detection is only carried out every *N-th* frame, with a value of $N = 15$ adjusted to the video used. While object tracing is performed between object detection breaks. The combination method is used to lighten the computational load but still gives good results [41]. The implementation of this method for the detection of poultry objects resulted in a Multi-Object Tracking Precision (MOTP) score of 60.4% with the results in the form of a graph of the movement and period of the object being detected.

3 Method

As a continuation of [24], this study uses the same object detection method, SSD, as a reference for assessment. The difference is in the object association method using the SORT algorithm and the object detection and tracking method is no longer carried out simultaneously. The scheme of this research is as shown in Fig.1.

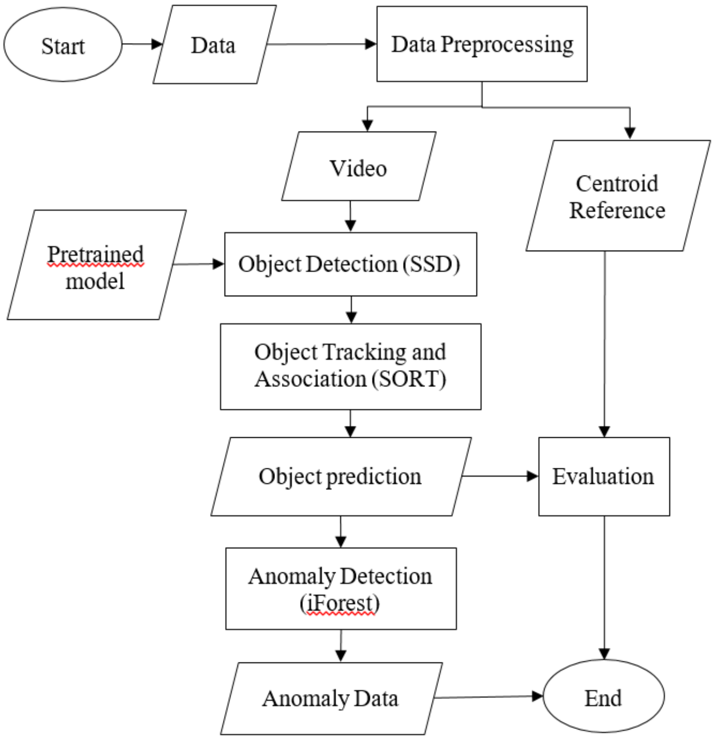


Fig. 1 Research steps

3.1. Dataset

This study uses two video datasets. The first dataset (data A) is the video data used in the research of [24] as a reference for the assessment of the algorithm, while the second dataset (data B) is video data taken using a Raspberry Pi device and a camera in the Field Section of Block B Poultry Unit, Faculty of Animal Husbandry, IPB University from 18 August to 13 September 202. The device was placed on four cages that were connected by the internet network. The position of the device and the top view of the chicken coop are shown in Fig. 2.

(a) camera position (b) top view

Fig. 2 Capturing device position

Table 1: Initial Video Specification

Data	Size(px)	FPS	Duration(seconds)
A	1280 x 720	30	55
B	1280 x 720	25	179

3.2. Data Preprocessing

Video data preprocessing is performed by changing the frame rate to 15 FPS. This is done to equalize the frame rate with previous research, besides that the lower frame rate also speeds up the object detection process and saves the memory of the device used. Data A and B are then extracted into frames. The next stage is to label objects in the frame in the first 30 seconds for algorithm evaluation purposes. An example of a frame from data A and B can be seen in Fig. 3.

3.3. Object Detection

Object recognition in data A and B is done using the Single Shot MultiBox Detector (SSD) algorithm. The basic network used in this research is the MobileNetwork (MobileNet) network. This network was chosen because it has a lower level of complexity than the VGG network [15]. The parameters used are the Intersection-over-Union (IOU) limit value of 0.6, confidence 0.6, and momentum 0.9. The detection model used at this stage is a pre-trained model from [24] which has been trained using the COCO dataset. This model is used as a reference for evaluation of previous research and accelerates the research process.

(a) (b)

Fig. 3 Examples of (a) data A and (b) data B

3.4. Object Tracking

In contrast to previous studies [24] that used Euclidean distance and centroid tracking algorithms, in this study the object identity association was carried out using the Simple Online and Realtime Tracking (SORT) algorithm which was built using the Kalman filter and the Hungarian algorithm [39].

3.5. Anomaly Detection

The research of [43] found that the small variation of the coordinates (x,y) and the object's velocity approaching the value 0 indicated the sitting behavior of the chicken object. The data generated from the object tracking stage is the displacement and time of the object's movement. In this study, anomalies are defined as data on the movement of objects with the longest idle time among other objects in the vicinity. The object idle time is calculated by subtracting the last time by the first time the object was detected. The longer the object is detected, the more likely it is that the object is stationary because a stationary object is easier to detect when compared to a moving object. If an object is detected as an anomaly then the object should be examined further.

Anomaly detection is carried out using the isolation forest (iForest) algorithm. The first step that will be carried out is the construction of the iForest model using the *sklearn* module, the default estimator parameter = 100 and contamination = 0.1. The *estimator* is used to determine the amount of data which is trained by the specified sample while the *contamination* value is used to determine the proportion of anomalous data in the data. The results of this stage are then processed so that they can display anomaly objects in the frame.

4 Results and Discussion

4.1. Data Preprocessing Results

This research was conducted using the hardware environment as follows: the processor is Intel®CORE™ i9 that is equipped with NVIDIA GeForce GTX 1650

GDDR5, 16 GB DDR4 RAM, 256 GB SSD and 1 TB SATA HDD. For programming, we use PyCharm Community Edition 2019.3.1 and Python 3.6.8. Video data A and B are converted using VSDC Free Video Editor so that it has a frame rate of 15 FPS. The results of data frame extraction from data A and B with a total of 448 frames each were then labeled using the Visual Object Tagging Tool (VOTT). Labeling is done on every frame and every chicken object seen in the video. An example results of the labeling can be seen in Fig. 4.

(a) (b)

Fig. 4 The (a) Source image and (b) Results of labeling process

4.2. Object detection results

Detection of objects using the SSD algorithm is carried out on data A and B. The detection model used has been adapted to research needs by selecting the 'bird' class with label 16, so that the detection model only stores the results of object detection that have a hypothesis class label 16. The $skip\_frame = 0$ parameter is used so that the detection model does not skip the target frame. From data A, the hypothetical data is generated for objects detected with a total of 1285 objects detected in 463 frames. Meanwhile, from data B, the hypothetical data of objects were detected with a total of 8717 objects in 4499 frames. The results of the detection stage are frames and files with Comma Separated Values (CSV) format with timestamp, runtime, frame, x1, y2, x2, y2, and score columns as shown in Table 2.

4.3. Object Tracking Results

The object detection data in the form of a CSV file is converted into JavaScript Object Notation (JSON) format. JSON is a data serialization format that is increasingly being used for data exchange [43].

From data A, the object has been traced 1016 times with the number of identities being 45 objects. Meanwhile, from data B, 7000 objects were searched with a total of 270 identities. Figure 10 shows the object search results in the form of a JPEG file. In addition to JPEG files, search results are issued in CSV format with frames, x1, y1, x2, y2, scores, id, date, and runtime columns as shown in Table 3. An example of a plot of the movement of data object A can be seen in Fig. 5.

Table 2: Object detection result

Timestamp	*Runtime*	*Frame*	x1	y1	x2	y2	*Score*
59:07.9	2.1875	1	630	222	746	385	0.957381
59:08.0	2.2187	1	572	551	724	707	0.896619
59:08.3	3.3281	2	569	554	721	712	0.983038
59:08.3	3.3593	2	631	222	747	386	0.973102
59:08.3	3.4062	2	819	437	972	550	0.628403

Table 3: Object tracking result

Frame	x1	y1	x2	y2	*Scores*	ID	*Date*	*Runtime*
/00004,0_mot,jpg	565	553	706	711	0,753561	195	10:38.4	125,8488
/00004,0_mot,jpg	631	223	747	383	0,753561	194	10:38.4	125,8494
/00005,0_mot,jpg	827	436	971	549	0,645254	196	10:38.5	125,9811
/00005,0_mot,jpg	567	554	703	710	0,645254	195	10:38.5	125,9817
/00005,0_mot,jpg	631	223	747	383	0,645254	194	10:38.5	125,9821

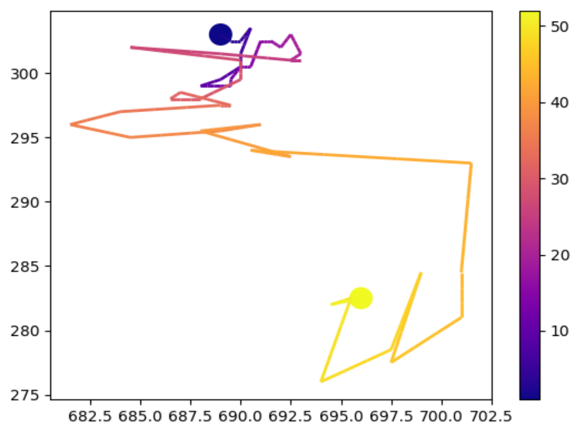


Fig. 5 The movement plot

The *frame* column contains the path of the frame when the object is detected, (*x1*, *y1*) is the lower left coordinate of the bounding box, (*x2*, *y2*) is the upper right coordinate of the bounding box, *scores* column contains the confidence value of the object, the *ID* column contains the identity of the object, the *date* column contains the date the search was performed, and the *runtime* column contains the time elapsed from the program running until the object was detected. The *runtime* column can be used to determine the frame [24] and the period of the detected object, which is useful in the analysis of broiler behavior anomalies.

4.4. Object tracking evaluation

Evaluation of object tracking results is done by calculating the value of Multi Object Tracking Precision (MOTP) [39]. MOTP was chosen as a comparison value between search results in previous studies with a MOTP score of 60.4% [24]. The step taken is to adjust the search result data so that it can be evaluated. Data adjustment is performed by calculating the value of the middle point of the bounding box (centroid) based on the coordinates of the bounding box of the object search results. The object's midpoint data can be seen in Table 4, where (cX, cY) are the coordinates of the bounding box's midpoint.

Table 4: Centroid data

Index	cX	cY
0	635.5	632.0
1	689.0	303.0
2	899.0	492.5
3	635.0	632.0
4	689.0	303.0

The next step is to calculate the distance between the midpoint of the object search results (hypothesis) and the midpoint of the reference data (actual). The reference data used is the data from object labeling at the data preparation stage. The distance between the midpoints is calculated using the Euclidean distance equation. The distance data between the midpoints can be seen in Table 5.

Table 5: Centroid distance data

Index	Δx	Δy
0	1.0	1.0
1	242.5	223.5
2	41.5	330.5
3	246.0	130.0
4	224.0	343.0

The distance data between the midpoints is then used as the basis for calculating the MOTP score. This study uses a limit value (threshold) for $\Delta x <= 118$ and $\Delta y <= 95$. This limit value is chosen because it is the average value of the width and height of the bounding box obtained using (6) and (7). The MOTP scores obtained in this study are 87.64% for data A and 85.67% for data B.

$$threshold_x = mean(\Delta x_0, \dots, \Delta x_n) \quad (6)$$

$$threshold_y = mean(\Delta y_0, \dots, \Delta y_n) \quad (7)$$

4.5. Motion behavior anomaly detection

The chicken movement anomaly in this study is the detected chicken object which has the longest detected time value, this is because the chicken object in a stationary position will have a more stable image and be easily detected by object recognition algorithms. Mathematically, the anomalous object of chicken movement is formulated in (8) with object identities from i to j and t as the time the object is detected.

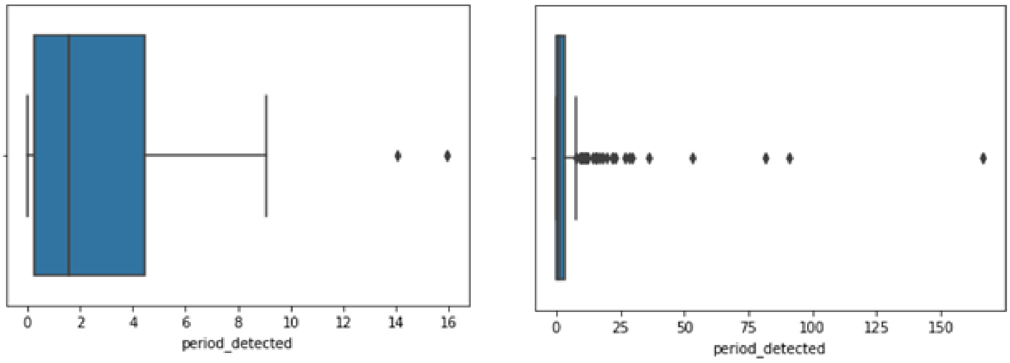
$$anomaly = arg\ max_{\ i,j}(\Delta t_i, \dots ,\ \Delta t_j) \quad (8)$$

Anomaly detection of movement behavior using the iForest algorithm is carried out by calculating the period for each object detected, so that data is generated as in Table 6 with the *ID* column containing the identity of the object, the *first_occurrence* column containing the time the object was first detected, *last_occurence* containing the time the object was last detected, and *period_detected* which contains the period the object was detected during the search. From this data, a visual boxplot can be made to determine the shape of the data and the existing outliers.

Table 6: Detected object period

ID	*first_occurrence*	*last_occurrence*	*period_detected (s)*
194	125.8494	132.6994	6.85005
195	125.8488	127.9487	2.09987
196	125.9811	126.3979	0.41681
198	126.9448	142.8943	15.9495
205	131.1958	131.4465	0.25067

Fig. 6 shows the existence of a period outlier value which is much higher than the other data. It is assumed that the object can be detected for a long time because it does not change position. Then by using the default estimator of 100 and contamination 0.1, anomaly detection is carried out using iForest. From the results obtained, it is known that there are anomalous objects marked with *anomaly_score* = -1 as shown in Table 7. The score can be calculated using (5) where E(h(x)) is the average of h(x) [25]. The identity association that is carried out still causes a change in identity for the same object, to find out whether the identity of the object refers to the same object can be seen from the identities of objects that have the same coordinates in the object search data.



(a) (b)

Fig. 6 Boxplot of detected object period of (a) data A and (b) data B

Table 7: Some of anomaly detection result

ID	*First_occ*	*Last_occ*	*Period_detected (s)*	*Scores*	*Anomaly_score*
194	125,8494	132,6994	6,850057	-0,0705	-1
195	125,8488	127,9487	2,099879	0,042162	1
196	125,9811	126,3979	0,41681	0,056043	1
198	126,9448	142,8943	15,94951	-0,27256	-1
205	131,1958	131,4465	0,250678	0,103237	1
208	132,9402	136,5213	3,581063	0,015165	1
209	133,3243	133,3243	0	0,0773	1
217	136,2771	138,7751	2,498021	0,021685	1
219	137,4681	144,8621	7,393979	-0,03217	-1

V. CONCLUSION

In this research a new dataset of chicken behavior has been developed, which is very valuable for further research of chicken behavior. Moreover, using the Simple Online and Real-time Tracking (SORT) algorithm and the Single-Shot Multi-Box Detector (SSD) object detection algorithm the MOTP score increase from 60.4% to 87.64%. In addition, the object detection period can also be used as an indication of an anomaly in the object's movement behavior. The association of the tracked objects with the iForest algorithm is successfully recognize the anomia of object movement behavior.

Our near future work is to implement the model in a real-time remote visual coop monitoring system to warn the staff if there is any unexpected condition in the cage. Moreover, the implementation of the model can be integrated into a complete surveillance system that includes environmental data such as temperature, humidity, ammonia level and noise level.

ACKNOWLEDGEMENTS
We express our gratitude to The Indonesian Minister of Research and Technology/National Research and Innovation Agency through IPB University for funding this research by the contract No. 3637/IT3.L1/PT.01.03/P/B/2022.

References

[1] BPS. 2020. Statistik Perusahaan Peternakan Unggas 2019.

[2] Soesanto IRH, Sudaryani T, Santosa H. 2011. *Panduan Lengkap Ayam*. Jakarta: Penebar Swadaya.

[3] Putri ED, Cepriadi C, Restuhadi F. 2020. Analysis of Production Efficiency Broiler Chicken Farm on Pattern Partnership of Contract Farming in Kampar District. *J Agribus Community Empower*. 3(1):60–70. doi:10.32530/jace.v3i1.94.

[4] Susanti S, Hidayati NA, Afriyansyah B. 2018. Penerapan Kesejahteraan Hewan Pada Beberapa Peternakan Ayam Ras Pedaging (Studi Kasus Di Kabupaten Bangka. *EKOTONIA J Penelit Biol Bot Zool dan Mikrobiol*. 2(1):49–57. doi:10.33019/ekotonia.v2i1.468.

[5] Bartussek H. 1999. Bartussek_1999_a review of the animal needs index.pdf. 61:179–192.

[6] Buijs S, Ampe B, Tuyttens FAM. 2017. Sensitivity of the Welfare Quality® broiler chicken protocol to differences between intensively reared indoor flocks: which factors explain overall classification? *Animal*. 11(2):244–253. doi:10.1017/S1751731116001476.

[7] Welfare Quality 2009. The Welfare Quality® assessment protocol for poultry (broilers, laying hens). The Welfare Quality® Consortium, Lelystad, The Netherlands.

[8] Farm Animal Welfare Council: Five Freedom. Available online: https://webarchive.nationalarchives.gov.uk/ukgwa/20121010012427/http://www.fawc.org.uk/freedoms.htm (accessed on 08 July 2022).

[9] Bessei W. 2006. Welfare of broilers: A review. Worlds Poult Sci J. 62(3). doi:10.1079/WPS2005108

[10] Berg C. 1998. Footpad Dermatitis in Broilers and Turkeys--prevalence, risk factors and prevention. Swedish University of Agricultural Science.

[11] Zhang Y, Wang C, Wang X, Zeng W, Liu W. 2020. FairMOT: On the Fairness of Detection and Re-Identification in Multiple Object Tracking. http://arxiv.org/abs/2004.01888.

[12] Angah O, Chen AY. 2020. Tracking multiple construction workers through deep learning and the gradient based method with re-matching based on multi-object tracking accuracy. *Autom Constr*. 119 November. doi:10.1016/j.autcon.2020.103308.

[13] Kraus M. 2020. Multi-Object Tracking in Aerial and Satellite Imagery. Technical University of Munich. https://elib.dlr.de/138070/1/ masterthesis_maximilian_kraus_final.pdf.

[14] Khan G, Tariq Z, Khan MUG. 2019. Multi-person tracking based on faster R-CNN and deep appearance features. *Vis Object Track with Deep Neural Networks*. https://www.intechopen.com/books/visual-object-tracking-with-deep-neural-networks/multi-person-tracking-based-on-faster-r-cnn-and-deep-appearance-features.

[15] Zhao ZQ, Zheng P, Xu ST, Wu X. 2019. Object Detection with Deep Learning: A Review. *IEEE Trans Neural Networks Learn Syst*. 30(11):3212–3232. doi:10.1109/TNNLS.2018.2876865.

[16] Liu W, Anguelov D, Erhan D, Szegedy C, Reed S, Fu CY, Berg AC. 2016. SSD: Single shot multibox detector. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)*. 9905 LNCS:21–37. doi:10.1007/978-3-319-46448-0_2.

[17] Chen L, Zhang Z, Peng L. 2018. Fast single shot multibox detector and its application on vehicle counting system. *IET Intell Transp Syst*. 12(10):1406–1413. doi:10.1049/iet-its.2018.5005.

[18] Miao X, Liu X, Chen J, Zhuang S, Fan J, Jiang H. 2019. Insulator detection in aerial images for transmission line inspection using single shot multibox detector. *IEEE Access*. 7:9945–9956. doi:10.1109/ACCESS.2019.2891123.

[19] Li Y, Hu W, Dong H, Zhang X. 2019. Building damage detection from post-event aerial imagery using single shot multibox detector. *Appl Sci*. 9(6). doi:10.3390/app9061128.

[20] Yang F, Chen H, Li J, Li F, Wang L, Yan X. 2019. Single Shot Multibox Detector with Kalman Filter for Online Pedestrian Detection in Video. *IEEE Access*. 7:15478–15488. doi:10.1109/ACCESS.2019.2895376.

[21] Liu P, Li X, Cui H, Li S, Yuan Y. 2019. Hand Gesture Recognition Based on Single-Shot Multibox Detector Deep Learning. *Mob Inf Syst*. 2019:25–28. doi:10.1155/2019/3410348.

[22] Huang X, Hu Z, Wang X, Yang X, Zhang J, Shi D. 2019. An improved single shot multibox detector method applied in body condition score for dairy cows. *Animals*. 9(7). doi:10.3390/ani9070470.

[23] Nagrath P, Jain R, Madan A, Arora R, Kataria P, Hemanth J. 2021. SSDMNV2: A real time DNN-based face mask detection system using single shot multibox detector and MobileNetV2. *Sustain Cities Soc*. 66. doi:10.1016/j.scs.2020.102692.

[24] Khairunissa J, Wahjuni S, Soesanto IRH, Wulandari W. 2021. Detecting Poultry Movement for Poultry Behavioral Analysis using the Multi-Object Tracking (MOT) Algorithm. *Proc 8th Int Conf Comput Commun Eng ICCCE 2021*.

[25] Liu FT, Ting KM, Zhou Z-H. 2008. Isolation Forest. In: *2008 Eighth IEEE International Conference on Data Mining*. IEEE. 413–422.

[26] Ounacer S, El Bour HA, Oubrahim Y, Ghoumari MY, Azzouazi M. Using Isolation Forest in anomaly detection: the case of credit card transactions. Periodicals of Engineering and Natural Sciences (PEN). 2018 Nov 24;6(2):394-400.

[27] Broom DM. Indicators of poor welfare. British veterinary journal. 1986 Nov 1;142(6):524-6.

[28] Broom DM. Animal welfare: concepts and measurement. Journal of animal science. 1991 Oct 1;69(10):4167-75.

[29] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556. 2014 Sep 4.

[30] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. Communications of the ACM. 2017 May 24;60(6):84-90.

[31] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. InProceedings of the IEEE conference on computer vision and pattern recognition 2015 (pp. 1-9).

[32] He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. InProceedings of the IEEE conference on computer vision and pattern recognition 2016 (pp. 770-778).

[33] Rezatofighi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S. 2019. Generalized intersection over union: A metric and a loss for bounding box regression. *Proc IEEE Comput Soc Conf Comput Vis Pattern Recognit*. 2019-June:658–666. doi:10.1109/CVPR.2019.00075.

[34] Ciaparrone G, Sánchez FL, Tabik S, Troiano L, Tagliaferri R, Herrera F. Deep learning in video multi-object tracking: A survey. Neurocomputing. 2020 Mar 14;381:61-88.

[35] Luo W, Xing J, Milan A, Zhang X, Liu W, Kim TK. Multiple object tracking: A literature review. Artificial Intelligence. 2021 Apr 1;293:103448.

[36] Chen L, Ai H, Zhuang Z, Shang C. Real-time multiple people tracking with deeply learned candidate selection and person re-identification. In2018 IEEE international conference on multimedia and expo (ICME) 2018 Jul 23 (pp. 1-6). IEEE.

[37] Jiang M, Rao Y, Zhang J, Shen Y. Automatic behavior recognition of group-housed goats using deep learning. Computers and Electronics in Agriculture. 2020 Oct 1;177:105706.

[38] Bernardin K, Stiefelhagen R. 2008. Evaluating multiple object tracking performance: The CLEAR MOT metrics. *Eurasip J Image Video Process*. 2008. doi:10.1155/2008/246309.

[39] Bewley A, Ge Z, Lionel O, Ramos F, Upcroft B. 2016. Simple Online and Realtime Tracking. *IEEE Int Conf Image Process*., siap terbit. https://arxiv.org/abs/1602.00763.

[40] Nascimento JC, Abrantes AJ, Marques JS. 1999. Algorithm for centroid-based tracking of moving objects. *ICASSP, IEEE Int Conf Acoust Speech Signal Process - Proc*. 6(1):3305–3308. doi:10.1109/icassp.1999.757548.

[41] Rosebrock A. 2018. OpenCV People Counter. *Pyimagesearch.*

[42] Aydin A, Bahr C, Berckmans D. Automatic classification of measures of lying to assess the lameness of broilers. Animal welfare. 2015 Aug 1;24(3):335-43.

[43] Severance C. Discovering JavaScript Object Notation. Computer. 2012 Apr 5;45(4):6-8.