Chicken Object Detection Using Faster Region-based Convolution Neural Networks

D.K. Gunabhiram

Abstract :

This paper presents deeper learning to identify different breads of poultry chickens under the context of poultry using a Faster Region-based Convolutional Neural Network (Faster R-CNN) with sample images totaling 605 images obtained from poultry houses. Three architectures of convolution were tested and implemented: ResNet-50, ResNet-101, and ResNeXt-101 for the best architecture for real-time chicken detection. In the context of ResNet-50, our results have demonstrated that we achieved higher success in Faster R-CNN. On average for your perusal, it has been achieved that the average precision at 50% is 96.453%, and at 75% it is 62.074%. Despite having more layers of convolution, the real-time processing time was faster in ResNet50 than in ResNet-101; thus, ResNet-50 was the most suitable architecture to be deployed for practical usage in poultry management systems. In this work, it is possible to present preliminary evidence of the effectiveness of faster R-CNN with ResNet-50 for further monitoring of cattle with high precision and speed.

Introduction :

The traditional means of counting the number of chickens manually in a poultry farm is very time consuming and involves lot of efforts plus it can easily incorporate the mistake of counting. Nevertheless, chicken object detection contributes highly to automation of counting and tracking in real-time on the chicken mass, through deep learning models such as Faster R-CNN with ResNet-50 through poultry farming. The real time manual supervision logged can be done apart from for the precise supervising of the poultry house climate by farmers with help of Faster R-CNN with ResNet-50, because this methodology has been proved to personal a very high level of accuracy and speed of processing. High accurate skin detection particularly at the 50% and 75% levels guarantee flawless performance regardless of the environment density. The ability of smart monitoring to real-time identification makes it possible to note an issue like overcrowding that may affect the wellbeing and productivity of chickens. Last but not least, this approach reduces the cost of labor while at the same time promoting effective management practices of poultry that enhance better health of the poultry and enhance operational efficiency in the poultry farming industries.

Computer vision is dominant in changing the poultry farming system since it brings about automated, precise, and continuous monitoring of the chickens. That is, computer vision facilitates a way through which the machine is capable of seeing and is similar to vision in a biological sense. With the help of machine learning and deep learning, computer vision is capable of identifying, categorizing and then monitoring objects in images or video flows effectively. When used in a poultry context, the technology renowned as computer vision, with elements of deep learning by the use of Faster R-CNN: ResNet-50 effectively enables the enumeration and detection of chickens in a poultry house. Having computer vision, farmers don't have to rely on observation which may take a long time and can often be imprecise. However, by using automation technology, one is able to monitor factors, such as flock size, flock density, and movements over time and generate real-time feedback and alarms. This technology makes it easier for farmers to provide the right living conditions for their chickens, be able to detect sickness and simple backing behaviours at an early stage, or even have more control over the production of eggs. Therefore, computer vision in applications such as chicken object detection improves intelligent poultry farming that fosters ethical treatment of chickens and farm efficiency.

Motivation

More specifically, this paper aims at designing and assessing a robust automated vision system that will be able to detect and monitor Chicks, during the poultry farming process. The Faster R-CNN with ResNet-50 model must solve the problems, which poultry farmers experience when they try to count and monitor the chickens on-farm, using only their sight, in order to offer the poultry industry a fast real-time robust high-accuracy chicken detection system. In this study, the performance of different architectures of CNNs is compared with the state of the art architectures including ResNet-50, ResNet-101 and ResNeXt-101 to determine the most effective architecture for real-time application of this concept to poultry management, thereby affording the capability for improved operational effectiveness.

In this paper, we suggested the detection and calculation of Chicken objects in a poultry house with the data characteristics to be used in the form of a video using one frame in two seconds, alongside standard still images of chickens. Object detection and calculation were done via the Faster R-CNN method integrated with Convolutional Neural Network (CNN) architecture.

Data used in this study is derived from the Hayriyigit Kaggle Chicken Object Detection data set which contains 376 images of chickens in natural butchered poultry farm circumstances. For better model training and improved detection accuracy, we used Albumentation which helps in data augmentation for the selected transformations including contrast adjustment, flip, and rotation. These augmentations added more variety to the dataset, ramping up the quantity of images in the set up to 605, which would be beneficial in making the model less specific. We then converted the annotations from the XML format in which they were created using LabelImg, to the COCO format, to enhance the compatibility of the dataset and ease of training with Faster R-CNN.

Novelty:

This project is novel in its application of advanced deep learning architectures such as ResNet50, ResNet-101, and ResNeXt-101 in real-time detection of chickens directly from live CCTV feeds. Most studies rely on static, preprocessed images, meaning that the dynamism and unpredictability of real video feeds are some of the things that this project would be exploiting to create more realistic monitoring and analysis in poultry environments. This brings in an entirely different dimension to the field because this state-of-the-art architecture captures much more sophisticated features and video input variations. The use of ResNet and ResNeXt on this front is also quite innovative, involving attention mechanisms and multi-path feature extraction, which are very well-suited for dealing with complex detection tasks that could be found in realtime. This new approach not only widens the technical methodologies used for animal monitoring but also bridges a gap toward their deployment in operational settings.

Contribution :

The work further advances real-time chicken detection, a capability that differs from previous studies whose image inputs were static or pre-recorded. Control settings with limited variability are used in most prior research whereas our approach captures the real world dynamics of poultry environments, including changing lighting and continuous motion, making this source of real time data worthwhile for practical deployment on poultry farms, which need uninterrupted, real time monitoring. This work offers a practical and implementable approach toward the perpetual monitoring on site—a domain that previous research has not really explored-through the examination of live video feeds' fluctuating and uncertain character.

A significant advancement can be attributed to our implementation of sophisticated deep learning architectures, namely ResNet-101 and ResNeXt-101, which are comparatively innovative within the realm of animal detection. Other than previous studies, which usually resort to the traditional CNNs or basic ResNet models, our selection of advanced models introduces sophisticated attention mechanisms and richer capabilities in feature extraction. These have enabled us to achieve an IOU of about 96% across the models. This puts a higher accuracy benchmark on real-time chicken detection. The high performance achieved demonstrates the robustness and precision of our approach, hence making it highly suitable for real-world applications where precision plays a critical role.

Nevertheless, a key perspective from this comparative evaluation of our project across different models—ResNet-50, ResNet-101, and ResNeXt-101—is something other studies haven't addressed.

This analysis highlights strengths and weaknesses of each model in the specific context of live CCTV video, hence supporting futures research and practical implementations by providing evidence-based model selection recommendations. This comparative approach adds a layer of contribution particular to surveillance-heavy applications, as it begins to lay down standards for choosing optimal models for specific tasks in real-time, accuracy-driven environments. This established a new yardstick against which the evaluation and selection of models in poultry monitoring systems would be measured.

Literature :

1) Counting Broiler from Poultry House Video Using Faster Region-based Convolutional Neural Networks

In the paper titled "Counting Broiler from Poultry House Video Using Faster Region-based Convolutional Neural Networks," Azzahrawan and Djamal are interested in utilizing faster and more effective approaches for using Faster RCNN in counting accurately placed broilers into poultry houses. The authors are using Faster RCNN, one of the best object detection models; they have such a robust two-stage architecture proficiently incorporating object region proposals along with accurate object detection capabilities. This allows the author to direct attention to difficulties that countering densely populated moving objects often ensues in restrictive environments.

The Faster RCNN architecture works by first using a Region Proposal Network (RPN) to propose possible locations for objects in an image. The RPN evaluates the input image to yield feature maps that mark where objects are likely to be situated, which are then passed to the second stage for both classification and localization. Using a Faster RCNN model supported by ResNet-101 was appropriate for this specific application because this is a kind of convolutional neural network architecture that favors vast feature extraction. The depth of the ResNet-101 enables the model to capture obscure details, making it more suitable for applications such as distinguishing the birds from the background and other entities in the poultry environment. Moreover, the RPN in Faster RCNN utilizes anchor boxes of different sizes and aspect ratios to propose regions which are to be adjusted according to variations in broiler sizes and shapes, which is a quite crucial feature for real-time counting in crowded spaces.

The present study made use of a dataset from the Gunther Cox GitHub repository, supplemented by additional images of broilers downloaded from Google. After annotating the images, they were resized as would be needed for training. With the RPN, it went through all the images for proposing locations of objects. Meanwhile, the Fast RCNN classifier goes through these proposals to enhance their accuracy. Among the architectures, ResNet-101 was deemed the best to get the most precision in detecting individual broilers. During the test, the Faster RCNN capability for video capture was validated with frames taken every two seconds. This frame selection strategy minimized data processing requirements while maintaining accurate broiler counts.

The results indicate that the Faster RCNN model using ResNet-101 achieved state-of-the-art performance on average precisions of 92.8% at IoU 0.50 and 78.1% at IoU 0.75.

These presented relevant results, showing an improved performance over the ResNet-50 and ResNeXt-101 architectures, highlighting the efficiency of ResNet-101 in feature extraction but countering the computational costs associated with ResNeXt-101, which, although deeper, had more latency due to its enormous model size. The region proposal network used anchor boxes set at three dimensions- 32x32, 64x64, and 128x128-with varying aspect ratios of 1:1, 1:2, and 2:1 so needed to effectively capture the wide range of appearances and positions of broilers.

The exclusion of region proposal approaches that often incur slower computing times and efficiency in real-time applications is the advantage of the Faster RCNN framework discussed in the paper. Because the integration of region proposals is effected directly within the Faster RCNN pipeline, this model facilitates a smooth merger between object detection and proposal generation notably improving both the speed and accuracy of counting of broilers in video frames. In summary, the authors of this manuscript show that the Faster RCNN architecture, using ResNet-101 as its underlying backbone, very efficiently counts broilers in a poultry setup. The power in the generation and classification of regions allowed the model to account for occlusion, motion, and density clustering problems. This work provides an excellent foundation for future studies on real-time livestock monitoring systems, even in the design of frameworks that require less computing power for in situ animal counting and tracking by machine.

2) Performance Analysis of YOLOv8, RCNN, and SSD Object Detection Models for Precision Poultry Farming

The research article "Performance Analysis of YOLOv8, RCNN, and SSD Object Detection Models for Precision Poultry Farming" assesses different object detection models targeted at determining the chickens and miscellaneous items in poultry situations. Specific scrutiny is made in regard to the framework of the object—RCNN, with a main focus on Faster RCNN—as compared with other models, including YOLOv8 and SSD object detection approaches, toward precision poultry farming for the recognition of myriad items, including chickens, feeders, and drinkers in an agricultural setting.

Faster R-CNN is known for its two-phase detection approach, starting with a Region Proposal Network (RPN) that detects the regions likely to contain the objects, followed by a convolutional network improving these proposals to classify and localise the objects selected. This approach makes Faster R-CNN highly suited to scenarios where object locations change and those objects to be searched by the camera show complex shapes, which is the scenario in poultry farming. Implementing anchor boxes of varying sizes and aspect ratios enables the model to recognize objects within a variety of scales and geometries, thus helping to detect chickens of various sizes in densely populated agricultural environments.

The authors evaluated the scores of mAP (Mean Average Precision) and the speed of Faster RCNN on their experiments. The model achieved an mAP@0.5 score of 0.59, which, while lower than YOLOv8's score, was sufficient for the study's purpose. Faster RCNN's relatively lower mAP score in this context is attributed to its slower inference speed compared to the YOLO models, which are optimized for real-time applications. The authors observed that although the network actually delivers considerable precision on object localization, its constraints on processing speed do not make it more suitable for fast, real-time application in poultry farming as operational demands must be in tune with fluctuating conditions.

It is quite vital for the RPN in Faster RCNN in order to efficiently provide region proposals within each frame likely to contain objects, thereby reducing the processing burden for the classification stage that follows. In this study, the RPN used three anchor box sizes and aspect ratios to cover a range of object scales within the poultry imagery. While the RPN allows for high object detection accuracy by preemptively focusing on probable object areas, it also introduces some processing delays, which were observed to impact the model's performance relative to single-stage detectors like SSD and YOLOv8.

The Faster RCNN model shows excellent localization and chicken classification accuracy in dense environments, which is a plus point, but, in fact, the authors argue that its applicability for real-time poultry monitoring is restricted since it has relatively higher computational needs. Differently than YOLOv8, which was superior in cases whose speed matters the most, the Faster RCNN fits much better with applications that demand high accuracy rather than high-speed detection.

This particular characteristic of the Faster RCNN underlines its application in environments where regular oversight is not imperative, or in instances where the necessary hardware for intricate calculations can be sustained. The analysis of Faster RCNN shows the paper supports it as a valid alternative where detection speed can compromise on its accuracy. The two-stage, structured architecture, with RPN for region proposals it has focused on the previous stages, does face the complexity of backgrounds and sized variation of objects present in poultry farms. For the practical use, the authors support the efficiency of YOLOv8 as well as compared to other alternatives because of its holistic accuracy and real-time performance balance.

3) Chicken Image Segmentation via Multi-Scale Attention-Based Deep Convolutional Neural Network

In the research paper titled "Chicken Image Segmentation via Multi-Scale Attention-Based Deep Convolutional Neural Network," the authors focus on the segmentation of chickens in poultry habitats, and rather particularly, strive for high precisions in segmentation areas with highly dense populations. While this manuscript largely presents a new multi-scale attentionbased deep convolutional neural network model known as MSAnet, references to Faster RCNN are also provided as a benchmark for the detection and segmentation of objects related to poultry.

The effectiveness of Faster RCNN in object detection does come from its two-stage approach: region proposals combined with object classification and localization. The RPN in Faster RCNN identifies regions that are supposed to contain objects; this aspect is most useful in crowded and overlapping environments like a poultry cage. Under these circumstances, Faster RCNN's RPN allows for accurate identification of individual chickens against the challenges posed by occlusion and varying object scales.

The authors investigate the segmenting ability of Faster RCNN when compared with MSAnet; it is clear that although in some cases, maximum detection accuracy by Faster RCNN can be achieved, in object segmentation tasks, it fails to accurately match the overlapping or touching objects. As Faster RCNN use multi-stage processing to achieve high performance in object detection, this inaccuracy will cause boundaries and appears during those applications that require very accurate segmenting details. Although MSAnet is particularly designed to enhance segmentation tasks based on the multi-scale and attention-based approaches, the role of Faster RCNN as a standard for comparison shows its comparative advantages and disadvantages in effectively demarcating different objects within intricate scenes.

In the context of anchor boxes, Faster RCNN makes use of different sizes with corresponding aspect ratios to adapt to the differences in chicken size at different stages of development.

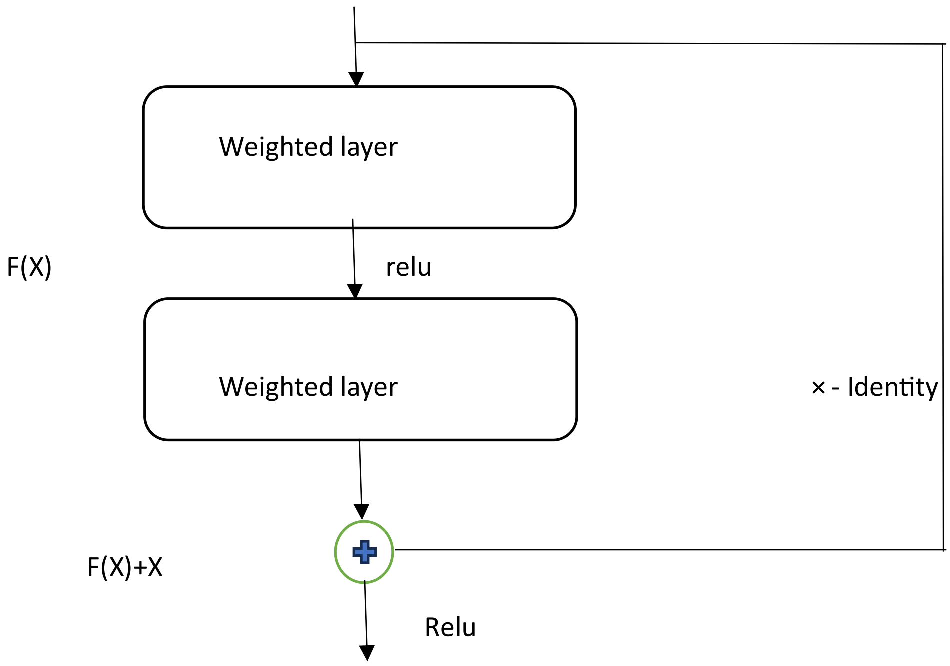
This allows for adaptability in chicken systems where those chickens are different sizes and especially do vary over time. The authors say, though, that this algorithmic method of Faster RCNN for generating region proposals adds overhead to processing and is thus slower at delivering segmentation outputs than MSAnet, which integrates segmentation features directly into its architecture. In this study, Faster RCNN functions as a benchmark for object detection, demonstrating the advantages associated with region proposal methodologies while also emphasizing the model's shortcomings in the area of segmentation. The research concludes that although Faster RCNN proves effective in detecting objects, the attention-driven methodology of MSAnet is more advantageous for pixel-level segmentation tasks within intricate poultry settings.

ResNet-50 CNN ARCHITECTURE Proposed Methodology :

ResNet-50 is a ground-breaking CNN that was introduced by Kaiming He and his colleagues at Microsoft Research Asia in mid-2015. ResNet is the abbreviation of the full name based on the network structure and can be mainly attributed to the residual blocks characteristic. This basic design comes from a residual learning framework in which it is possible to train very deep networks with several layers, this probably would have been impossible in the past because of the depth limitations of deep learning.

The inception of ResNet was driven by a critical observation in deep learning research: Cognitively, argumentations showed how the intuitive notion of how extra layers in the neuronal nexus of the neural networks would improve its efficiency was actually false. Classical theory would expect that more profound networks should always perform better and yet at the same they demonstrated, through experiments, that increasing layers just provided counterproductive results.

These skip connections eliminated the main difficulty known as the vanishing gradient, observed chiefly during the training phase of deep network architectures. Thus, ResNet made it possible to use backpropagation to train models with up to 152 layers and so improving the capacity of deep neural networks in the learning of input data representation. This change is considered to be a revolution in the progress of deep learning making the contribution to the evolution of way deeper and more powerful architectures.



Skip Connection

Three key components of ResNet

Convolutional Layers:

The network begins this by using convolutional layers on the input images. These layers are responsible for feature extraction: they multiply the input with different kernels to find out features such as edges, textures and patterns. The amount of output classes is often less than the amount of transformations made on input images, rarely it is equal. In this case, the convolutional layers move over the input data converting it to feature maps, which depict salient features of data sets.

LayerName	**OutputSize**	**50-Layer**	**101-Layer**
Convl	**112x112**	**7x7,64,stride2**	
		3x3maxpool, stride2	
Conv2_x	56x56	1x1,64 [3x3,64]x3 1x1,256	1x1,64 [3x3,64]x3 1x1,256
Conv3	28x28	1x1,128 [3x3,128]x4 1x1,512	1x1,128 [3x3,128]x4 1x1,512
Conv4_x	14x14	1x1,256 [3x3,256]x6 1x1,1024	1x1,256 [3x3,256]x23 1x1,1024

Conv5	7x7	1x1,512 [3x3,512]x3 1x1,2048	1x1,512 [3x3,512]x3 1x1,2048
	1x1	averagepool,1000-dfc,softmax	
	FLOPs	3.8x109	7.6x109

Residual Blocks:

Currently tunared linear unit or ReLU activation functions accompany two convolutional layers in each residual block. To further normalize the features extracted from the convolutional layer it is used. Surprisingly, the result in the second convolutional layer enhances the input of the residual block, starting with one more 56ReLU activation. This amplified data is then forwarded for further processing to the next block in the architecture.

Fully Connected Layer:

Intended to predict the output from the last residual block against some target output classes, the last layer of the architecture is fully connected. The total number of output classes is proportional to the number of neurons in it. In other words, ResNet is built with an integrated architecture comprising convolutional portions, residual parts, and fully connected portions.

The architecture of ResNet Layers

ResNet-50 Model Architecture

INPUT

Conv 7x7 s:2	
Batch Normalization	
Activation	
Maxpooling 3x3 s:2	
Conv [block]	
Identity [block] Conv [block]	**x2**

Identity [block] Conv [block]	**x3**
Identity [block] Conv [block]	**x5**
Identity [block]	**x2**
Average pooling	
Full conv	

↓

OUTPUT

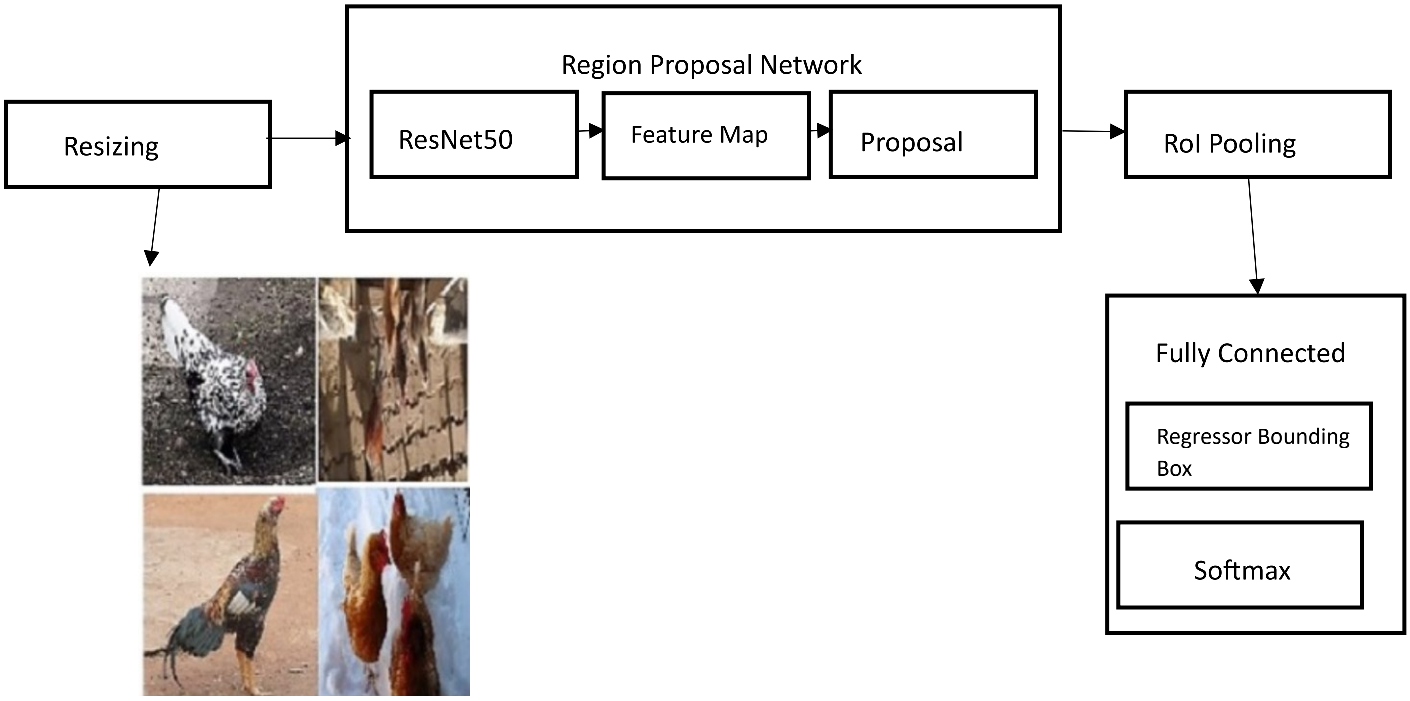
Advantages of the ResNet-50 over the Other Networks

There is a number of benefits related to the use of ResNet-50 among other networks. The major advantage is its capacity to train networks of a very high order, that is, networks with hundreds of layers. This is attributed to the fact that, the network architecture incorporates the use of residual blocks and skip connections through which information from initial depths of the network are retained.The other advantage of ResNet-50 used in this study is that it has been proven to give state-of-the-art results in various image-related applications including object detection, image classification and image segmentation..

Experimental Setup and Result:

Broiler counting system using Faster R-CNN

The proposed approach for broiler livestock detection utilizes Faster R-CNN with ResNet-50 as the CNN backbone. Initially, the input image goes through a resizing stage to adjust its dimensions, making it compatible with the CNN. The resized image is then passed to the Region Proposal Network (RPN), which generates proposals for potential objects in the form of predicted class scores and bounding box dimensions. These proposals are then processed by the Region of Interest (RoI) layer, which converts the RPN output matrix into a vector format suitable for the Fully Connected layers, where the proposals are evaluated.



DATASET :

The dataset for this study was drawn from the Hayriyigit Kaggle Chicken Object Detection dataset. It comprises 376 images of chickens taken in natural settings, as well as butchered poultry farms. We used the Albumentations library-the powerful library for data augmentation; its transformations also include adjusting the contrast, flipping, and rotating in contrast with the other libraries. These augmentations expanded the diversity of the dataset to 752 images, helping the model generalize and reduce the overfitting effect.

The annotations in XML format (LabelImg) were converted to the COCO format for better compatibility of datasets and easier training with the Faster R-CNN model. The training set uses 80% of the augmented dataset and the other 20% for testing purposes, thus allowing proper evaluation of the model.

Dataset (a) Hayriyigit Kaggle

Dataset (b) Augmented images using Albumentations

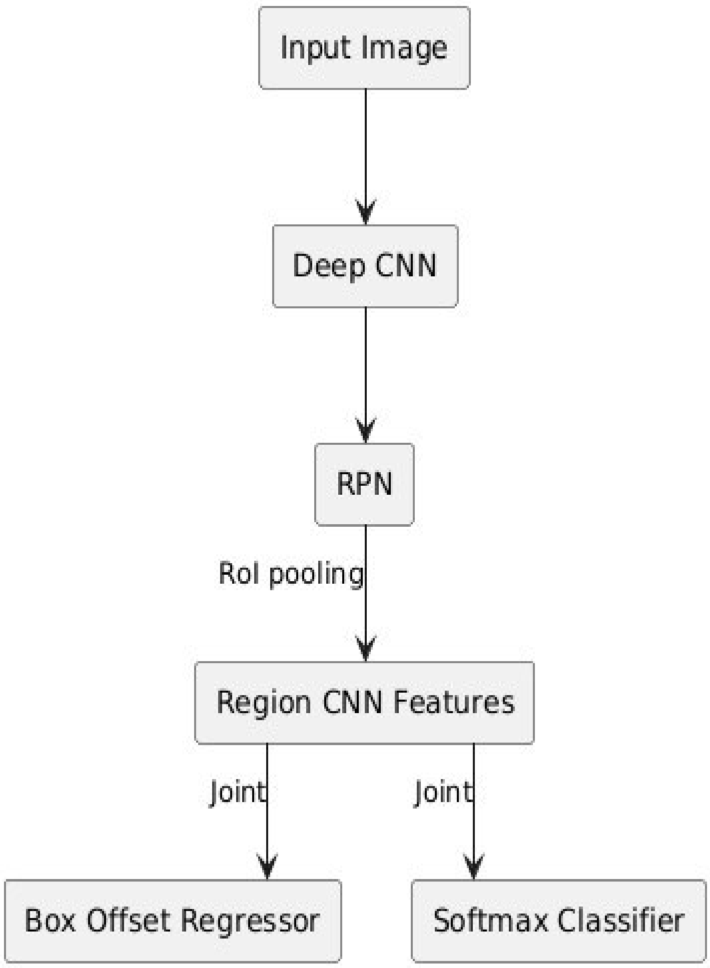
The dataset has not been annotated, which means the dataset cannot be used to carry out the training process on the Faster R-CNN model. So, all the chicken images need to be annotated. An application called to make sense annotating images. An example of annotating an image can be seen here

Annotating images using labelImg

Faster R-CNN with Detectron2 :

Detectron2 was built by Facebook AI Research (FAIR) to support rapid implementation and evaluation of novel computer vision research. In this paper, Detectron2 was providing the model that Faster R-CNN utilised; it is a platform by Facebook AI Research to give users many object detection models while streamlining and accelerating the creation of such models. The model was given and already trained with the COCO dataset. Then, Faster R-CNN would consist of two parts- the Region Proposal Network (RPN) and Fast R-CNN as the detector. RPN would have a portion of CNN to extract the features that would ultimately lead to the Feature Map.

This Feature Map is then used to propose the coordinates and class of detected objects. In this paper, the version of CNN architecture implemented is ResNet-50 which is utilized to provide a more shallow yet efficient substitute to ResNet-101. As the dimensions of the proposals that are generated by the RPN are different, ROI Pooling is done for uniformity among them. The pooling in the ROI Pooling gives equal and fixed number of proposals. In the stage of Fully Connected it is possible to calculate the likelihood for object classification along with refining the coordinates of the bounding box for providing as final outputs for object detection.



Faster R-CNN (Using PantUml)

Region Proposal Network (RPN) with ResNet50 Backbone

Creating Feature Maps with ResNet50:

The backbone model we use here is ResNet50, and this backbone model scans over the input image by means of a number of consecutive convolutions and pooling layers in order to deliver a feature map, on which the RPN can scan and propose potential locations of objects across the image.

Anchor Boxes: Anchor boxes are predetermined bounding boxes with varying sizes and aspect ratios, assisting the RPN to identify potential objects of various dimensions and shapes.

Anchor Sizes: Anchors with sizes between 32x32, 64x64, 128x128, 256x256, and 512x512 pixels in size are utilized by the RPN. The entire range allows for the detection of objects that start from being very small and can be large.

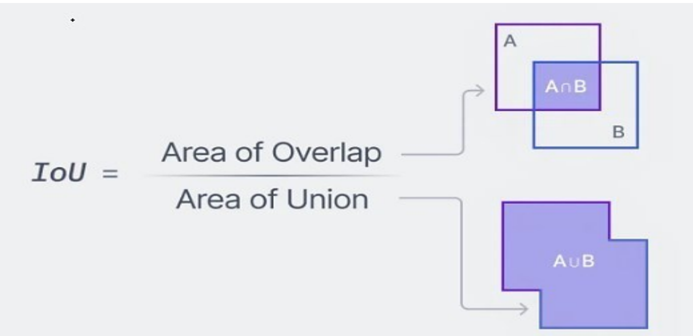
Aspect Ratios: The anchor boxes also comes in different aspect ratios: 1:1, square; 1:2, tall; and 2:1, wide. This is to help the RPN detect objects with different shapes.

At every location on the feature map, 15 anchor boxes are created. They cover a wide range of sizes and shapes.

Selection based on IoU: Anchors For each anchor, for the anchor box and the ground truth boxes IoU score is computed. It determines how much the two boxes overlap with each other.

Intersection over Union (IoU): IoU is calculated as:

$$\text{IoU} = \frac{\text{Area of Overlap}}{\text{Area of Union}}$$



where Area of Overlap is the region between the anchor and the ground truth box, and Area of Union is the total area covered by both the boxes. A better match of an anchor box with a ground truth box is obtained for each object for which the anchor box having the highest IoU is chosen, and that anchor is refined to enhance the precision in estimating the location as well as the size of the object.

Fast R-CNN :

In Faster R-CNN with ResNet50, the RPN directly generates object proposals, hence giving up the use of selective search methods. Then these proposals are pooled to a fixed size using Region of Interest (RoI) Pooling where each proposal is divided into a grid, and each cell's largest value is selected to standardize dimensions. Fast R-CNN then classifies these proposals and refines their bounding boxes for accurate object detection.

Results and Discussion :

In this experiment, the Faster R-CNN model was fine-tuned on the COCO dataset. Detectron2 is multicontext CNNs, which evaluated using different architectures like ResNet-50, ResNet101, and ResNeXt-101. Its effectiveness is examined using both images and video of broilers, using its application in a video to count the number of detected broilers in every frame, two seconds apart. This approach will be helpful to find out which CNN architecture can perform best in detecting broilers for still images and for video. The training process results on several CNN architectures used can be seen in Table 1.

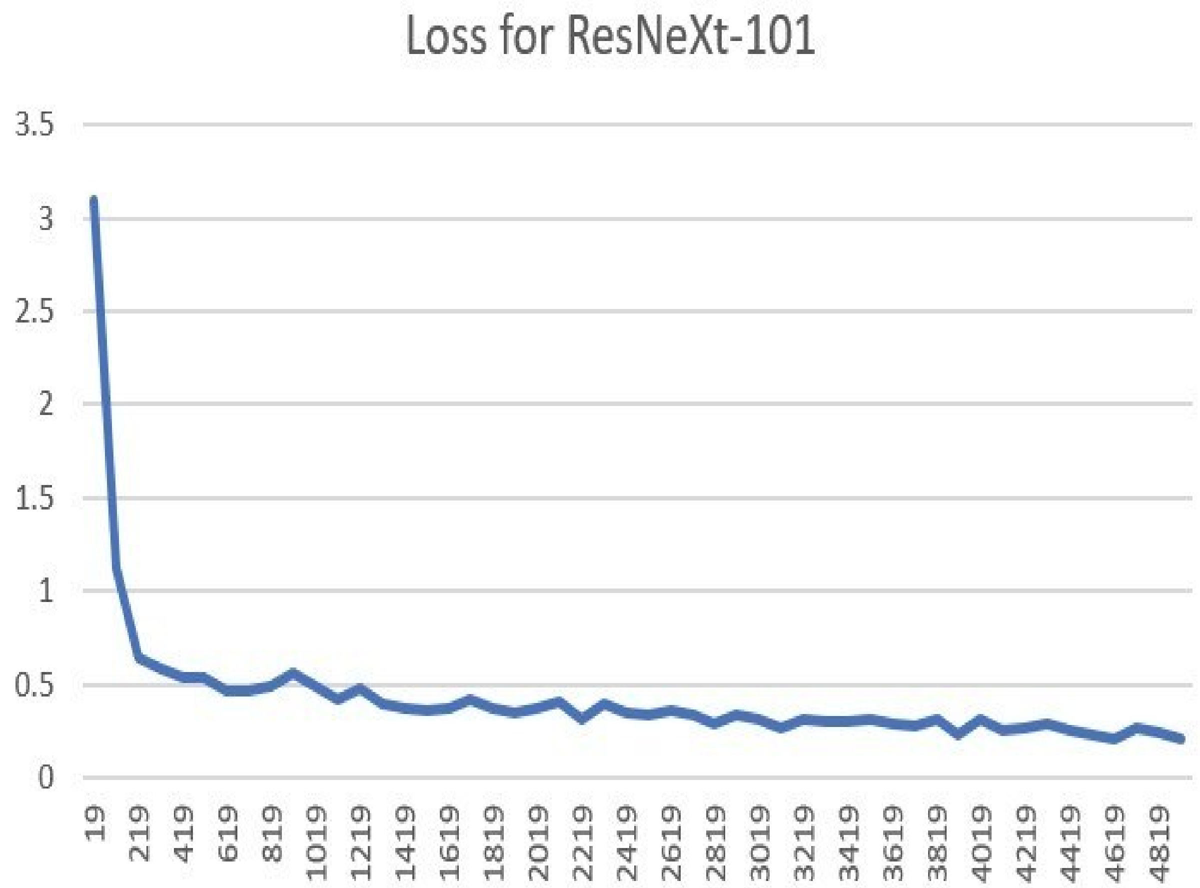
Table (1): EVALUATION PROCESS OF TRAINING MODEL

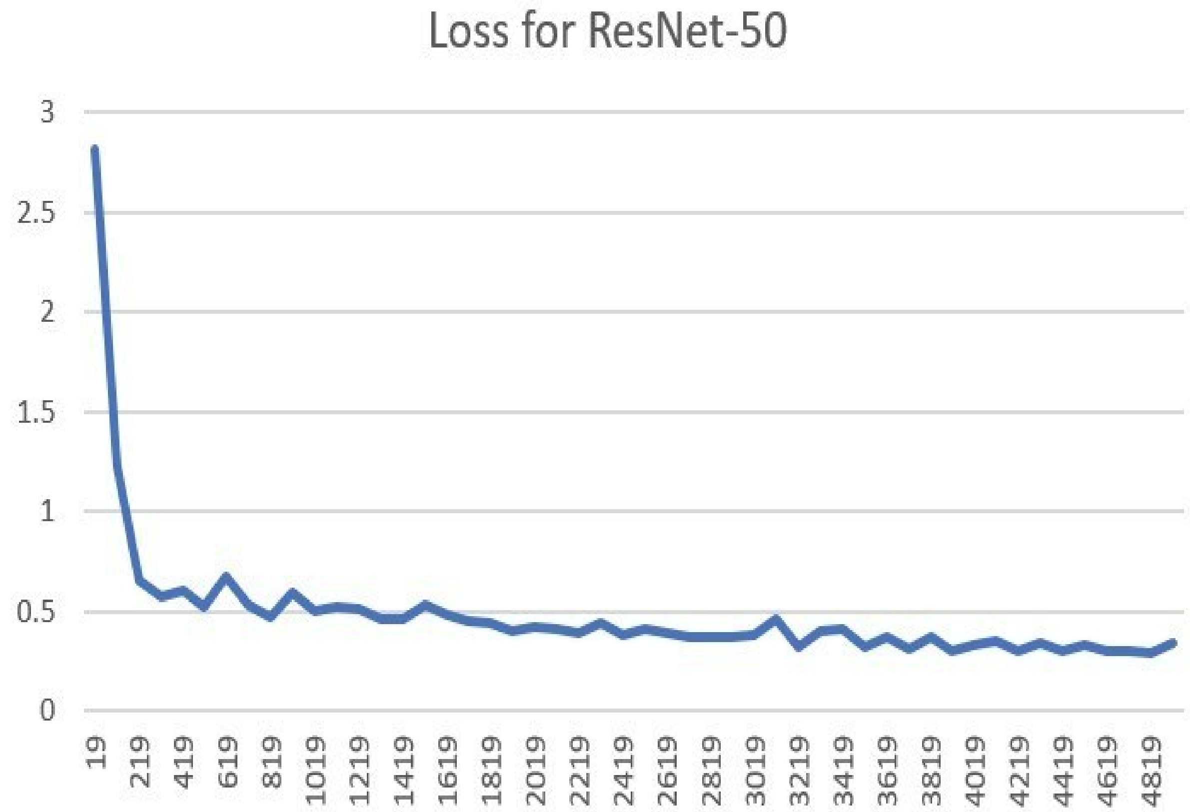
CNN Architecture	Average Precision IoU (%) - 0.50	Average Precision IoU (%) - 0.75	Total Loss	Size (MB)	Training Time
ResNeXt-101	96.267	61.796	0.2137	796.70	Slowest (85 min)
ResNet-50	**96.453**	**62.074**	**0.3200**	**314.93**	**Faster (24 min)**
ResNet-101	96.084	62.653	0.3027	460.81	Slower (34 min)

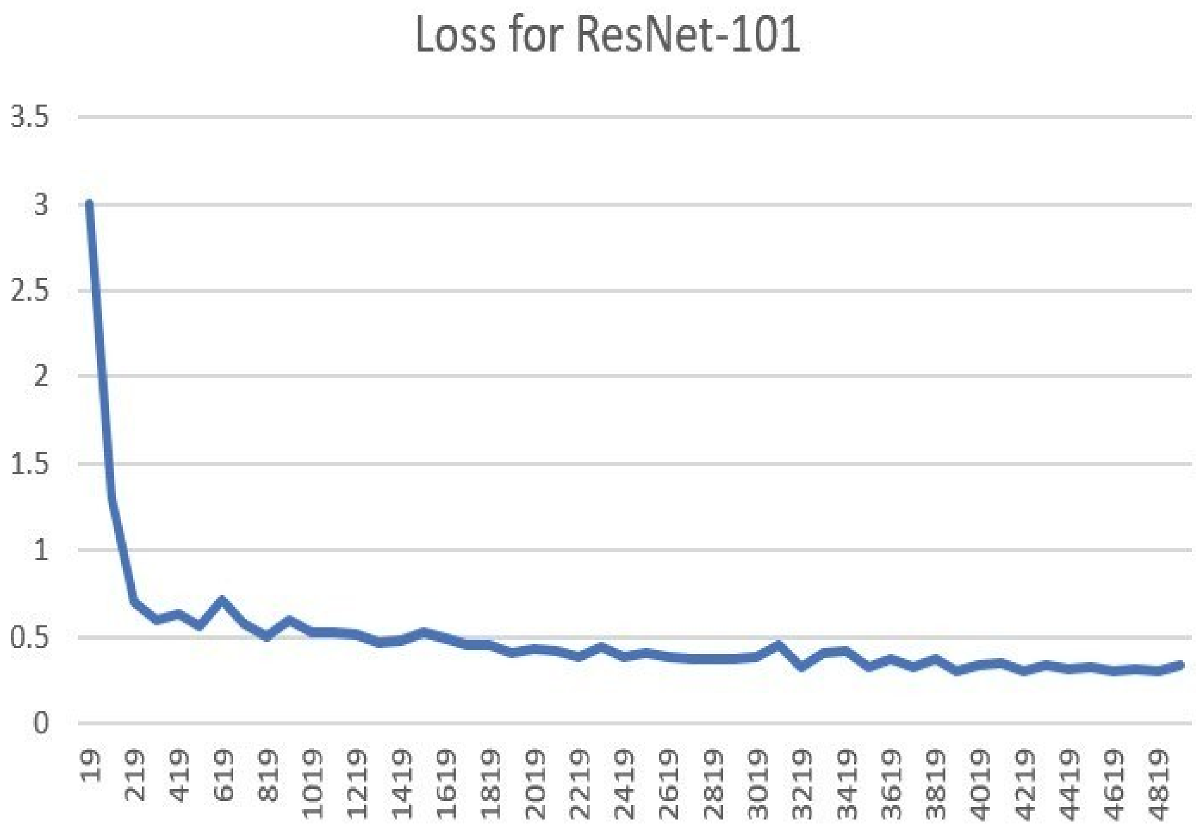
Based on table (1), the average precision with IoU was 50% and in 75% used ResNet-50, CNN Architecture had the highest accuracy. Also, the size of the model using ResNet-50 is smaller when compared to ResNet-101 and ResNeXt-101, with a difference of size 145.88 MB smaller than ResNet-101 and with a difference of size 481.77 MB smaller than ResNeXt-101

Inference from Training Loss Convergence for Different CNN Architectures

The graph of loss reduction with the number of iterations for ResNet-50, ResNet-101, and ResNeXt-101 indicates that ResNet-50 converges more rapidly. It reaches a smaller value of loss in fewer iterations due to a smaller architecture, hence it is efficient in using for tasks that require shorter times for training. ResNeXt-101, having a deeper structure, requires more time to converge but is able to get similar final values of loss since it can learn more complex features. The slowest one to converge is probably the complex architecture of ResNeXt-101; however, considering its complexity, it would probably take on better tasks in more demanding applications.




Loss for ResNet-50
3
2.5
2
1.5
1
0.5
0
19 219 419 619 819 1019 1219 1419 1619 1819 2019 2219 2419 2619 2819 3019 3219 3419 3619 3819 4019 4219 4419 4619 4819


Loss for ResNet-101
3.5
3
2.5
2
1.5
1
0.5
0
19 219 419 619 819 1019 1219 1419 1619 1819 2019 2219 2419 2619 2819 3019 3219 3419 3619 3819 4019 4219 4419 4619 4819

Comparision Table :

	Chicken Object Detection-Gunabhiram, Pranith, and Jaswanth	Counting Broiler from Poultry House IEEE 2022-Muhammad Reza Azzahrawan, Esmeralda Contessa Djamal	Performance Analysis of RCNN for Poultry Farming Vishnu Kumar Kaliappan, Manjusree S V, Gangadhar Baniekal Hiremath
CNN Architecture	**Average Precision IoU (%) - 0.50**	**Average Precision IoU (%) - 0.50**	**Average Precision IoU (%) - 0.50**
ResNeXt-101	**96.267**	90.60	93.71
ResNet-50	**96.453**	90.84	94.45
ResNet-101	**96.084**	92.82	92.08

This table summarizes the performance of different CNN architectures: ResNeXt-101, ResNet50 and ResNet-101 on the chicken data in terms of AP at an IoU of 0.50. Gunabhiram, Pranith, and Jaswanth study concludes that the highest AP 96.453% is reached for ResNet-50, followed by ResNeXt-101 96.267% and then ResNet-101 with 96.084%. In a 2022 IEEE report by Muhammad Reza Azzahrawan and Esmeralda Contessa Djamal, ResNet-101 ensured the highest AP to be 92.82%. ResNet-50 was close to that with an AP of 90.84% and then ResNeXt101 secured an AP of 90.60%.

Finally in the work by Vishnu Kumar Kaliappan, Manjusree S V, and Gangadhar Baniekal Hiremath, the highest AP of 94.45 was ensured by ResNet-50 and then 93.71% by ResNeXt101 followed by 92.08% by ResNet-101. Overall, ResNet-50 exhibits the least variability in consistently maintaining high performance across studies. ResNeXt-101 and ResNet-101 also have variable performance; at times, the former does better than the latter, though rarely, while the latter does better than the former.

Comparative with State-of-the-Art with others Work :

Compared to the state-of-the-art approaches, this research considered the performance of Faster R-CNN with three other variants of a backbone CNN architecture, namely ResNet-50, ResNet101, and ResNeXt-101, for real-time detection of chickens on a poultry farm. Here's how this research compares to other notable works:

ResNet-50 Performance:

It achieved an average precision of 96.453% at IoU 0.50 and, therefore, represents the best chicken-detecting model in consideration with practical accent on the speed and efficiency. Compared to ResNet-101 and ResNeXt-101, ResNet-50 showed a better trade-off concerning model size in precision and leads to faster inference times that is ideal for real-time monitoring in poultry farms. This model, in earlier studies like the 2022 study by Azzahrawan and Djamal, came together with Faster R-CNN to use ResNet-101 for counting broilers, reaching an AP of 92.8% at IoU 0.50. While ResNet-101 could capture details better because it had more depth, it consumed a lot of time to process, which was slow for real-time applications.

ResNeXt-101 vs ResNet-101:

Although ResNeXt-101 was good at capturing complex features, the model size was larger and required higher computational power than that of ResNet-50; thus, processing speed on realtime applications slowed down. Conclusively, while maintaining high precision, being more computationally efficient made ResNet-50 outstand from the rest, including ResNet-101 and ResNeXt-101 for deployment in real-world poultry monitoring applications. This test showed that ResNet-101 achieved an AP of 96.084% but was more time-consuming to process images compared to ResNet-50. So, although ResNet-101 is especially good for applications requiring maximum accuracy, the speed advantages of ResNet-50 make it a better option for continuous, real-time detection applications.

Efficiency of the Faster R-CNN Framework:

Further, the Faster R-CNN architecture, which has been used in this work also, is based on the approach seen in most of the recent works such as comparison studies of YOLO and SSD, intended for real-time applications in farming. However, two-stage process handling of Faster

R-CNN facilitates maintaining high precision in detection for objects since it makes region proposals first before classifying and localizing objects in those regions. While single-stage detectors such as YOLO are much faster than Faster R-CNN, the latter is still a good balance for any environment where accuracy matters.

Further Reach and Model Agility :

Findings from this research expand the scope of deep learning models in poultry management and make poultry farming more efficient. It establishes a benchmark for different performance metrics in chicken detection in this setup, which will guide future studies in choosing model architectures for similar applications. Summarily, the experiment conducted in this work sets a new benchmark standard for detecting objects from a chicken populace on poultry farms using ResNet-50 on top of the Faster R-CNN framework. This will be because the accuracies surpassed or closely approached those obtained from other state-of-the-art models including ResNet-101 and ResNeXt-101 while being made certain that the processing speeds remained faster. Thus the work shows that ResNet-50 is a powerful and efficient model choice for realtime poultry monitoring.

Results :

To test object detection, three test images were utilized, as depicted in Figure 6. For the new thresholds of 50% and 75% IoU, they have been tested. The process, therefore, culminates into the most effective settings for the precise identification of the objects. The ResNet-50 architecture, which proved efficient and fast with regards to accuracy performance in short time intervals, was used to test the object detection. These IoU thresholds balanced the speed and detection quality of the model. Thus, ResNet-50 was an optimal choice for real-time applications such as quick and accurate object detection in that environment.

The testing setup was the same as for broiler images, only with object detection on one frame every two seconds, and the evaluation was through the count of correctly detected broilers versus the count of wrongly detected objects. In this setup, Faster R-CNN ResNet-50 outperformed ResNet-101 and ResNeXt-101 architectures, as those had troubles with accuracy due to non-broiler objects being detected. The accuracy of ResNet-50 was better and followed the images more strictly than any other architectures; hence it was the best alternative for this task.

It correctly identified the chicken

Conclusion :

This study concludes that Faster R-CNN with ResNet-50 was the best-performing chicken detection model from real-time poultry environments. Among architectures tested, the one presented here achieved a good trade-off between accuracy and processing speed by attaining an average precision of 96.453% at IoU = 0.50, which makes it suitable for deployment in operational settings with real-time monitoring requirements. ResNet-50 has the advantage of being smaller and less computationally expensive than the deeper architectures of ResNet-101 and ResNeXt-101. Therefore, in real-time detection, it can aptly be used for practical purposes. These employ practical requirements where fast and reliable monitoring would imply a reduction in labor, better poultry welfare by providing continuous oversight, and overall productivity in poultry farms. In comparison, results suggest that the application of ResNet-50 in the Faster R-CNN framework achieves real-world precision needs and provides access to an available solution for farm automation and management needs of poultry farms where real-time insights are critical for maintaining optimal poultry health, efficiency, and operations.

Future Work :

Future work on this project could include further optimizations to make the model run faster while keeping high accuracy. This would then be actually more deployable for applications with real-time inference requirements. A possible direction is to look into other light-weight CNN architectures, such as MobileNet or EfficientNet, which are popular due to good performance with respect to lower computational requirements. Another way to make such a model more versatile and helpful to a larger set of needs in agricultural monitoring is to expand its application to other kinds of livestock or animals typically located in farmland. Other avenues of improvement could be the development of methods to enhance the robustness of the model, for example, when there's much occlusion or varying lighting conditions, with respect to difficult cases. An adaptive system that can continue to learn from new farm data can also improve model adaptability and long-term accuracy. Innovation can thus lead to much more efficient, accurate, and flexible systems for monitoring farm environments and managing livestock welfare.

References :

1) S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE transactions pattern analysis and machine intelligence, vol. 39, no.6,pp. 1137-1149, 2017.

2) Muhammad Reza Azzahrawan, Esmeralda Contessa Djamal and Jenderal Achmad Yani Cimahi, " Counting Broiler from Poultry House Video Using Faster Regionbased Convolutional Neural Networks"DOI: 10.1109/ISITIA56226.2022.9855276

3) X. Zhuang and T. Zhang, "Detection of sick broilers by digital image processing and deep learning," Biosystems Engineering, vol. 179, pp. 106–116, 2019.

4) M. B. Jensen, M. P. Philipsen, A. Møgelmose, T. B. Moeslund, and M.M. Trivedi, "Vision for Looking at Traffic Lights: Issues, Survey, and Perspectives," IEEE Transactions on Intelligent Transportation Systems, vol. 17, no. 7, pp. 1800–1815, 2016.

5) Y. Wu, A. Kirillov, F. Massa, W.-Y. Lo, and R. Girshick, "Detectron2." 2019.

6) S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 39, no. 6, pp. 1137–1149, 2017.

7) R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Region-Based Convolutional Networks for Accurate Object Detection and Segmentation," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 1, pp. 142–158, 2016.

8) R. Girshick, "Fast R-CNN," Proceedings of the IEEE International Conference on Computer Vision, vol. 2015 International Conference on Computer Vision, ICCV 2015, pp. 1440–1448, 2015.

9) H. Zhang, Y. Du, S. Ning, Y. Zhang, S. Yang, and C. Du, "Pedestrian Detection Method Based on Faster R-CNN," Proceedings - 13th International Conference on Computational Intelligence and Security, CIS 2017, vol. 2018-Janua, pp. 427–430, 2018.

10) R. I. H. Abushahma, M. A. M. Ali, O. I. Al-Sanjary, and N. M. Tahir, Region-based convolutional neural network as object detection in images," Proceeding - 2019 IEEE

7th Conference on Systems, Process and Control, ICSPC 2019, no. December, pp. 264–268, 2019.

11) X. Zhuang and T. Zhang, "Detection of sick broilers by digital image processing and deep learning," Biosystems Engineering, vol. 179, pp. 106–116, 2019,

12) B. Zhu, X. Wu, L. Yang, Y. Shen, and L. Wu, "Automatic detection of books based on Faster R-CNN," 2016 3rd International Conference on Digital Information Processing, Data Mining, and Wireless Communications, DIPDMWC 2016, pp. 8–12, 2016,