Dataset Creation for the Artificial Intelligence-based Video Analysis of Yellow Hungarian Chicken Flocks

Adathalmaz készítése sárga magyar tyúk állományok mesterséges intelligencia alapú videóelemzéséhez

Éva Rampasek[1*], Csaba Hajdu[2], Boldizsár Tüű Szabó[3] Károly Tempfli[4] and László Környei[5]

[1]*Multidisciplinary Doctoral School of Wittmann Antal Crop-, Animal- and Food Sciences; Faculty of Albert Kázmér Mosonmagyaróvár Széchenyi István University; rampasekeva@yahoo.com*

[2]*Department of Informatics, Faculty of Mechanical Engineering, Informatics and Electrical Engineering, Széchenyi István University; hajdu.csaba@ga.sze.hu*

[3]*Department of Informatics, Faculty of Mechanical Engineering, Informatics and Electrical Engineering, Széchenyi István University; tuu.szabo.boldizsar@sze.hu*

[4]*Department of Animal Science, Faculty of Albert Kázmér Mosonmagyaróvár Széchenyi István University; tempfli.karoly@sze.hu*

[5]*Department of Mathematics, Faculty of Mechanical Engineering, Informatics and Electrical Engineering, Széchenyi István University; laszlo.kornyei@math.sze.hu*

**Correspondence: rampasekeva@yahoo.com*

Abstract: A Precision Livestock Farming (PLF) system offers real-time insights into animal welfare and health, increasing production efficiency. Computer vision in the poultry industry enables automated monitoring of animals, providing a wealth of information without the need for human resources. The adoption of the "End the Cage Age" initiative by the European Commission will phase out caging of farm animals expectedly from 2027, impacting the laying hen sector, especially in Hungary where a significant portion of hens are currently raised in cage systems that are being replaced with cage-free systems requiring continuous monitoring of social behaviour. The initial step in implementing machine vision involves identifying hens and determining the number of individuals visible in images. The study involved automatic monitoring of various aspects such as behaviour, welfare, and health of Yellow Hungarian chickens at a breeding farm in Mosonmagyaróvár. Video recordings captured the activities of 49 hens in a controlled environment, focusing on detecting individual bird movements and predicting behaviours using advanced image processing techniques. The researchers utilized open bird datasets and annotation tools to create a specialized dataset for chicken behaviour analysis, showing promising results for future enhancements in detection and prediction capabilities.

Keywords: *Hens, YOLO, precision livestock, deep learning, object tracking, object detection, artificial intelligence, computer vision*

Összefoglalás: A precíziós állattenyésztési rendszerek (PLF) képesek valós időben folyamatos képet adni az állatok jólléti, valamint egészségi állapotáról, és egyben a termelés hatékonyságának növelését is biztosíthatják. A számítógépes látás az információs technológiák fejlődésének köszönhetően a baromfiiparban az állatok megfigyelését teszi lehetővé emberi erőforrás igénybevétele nélkül és számtalan információ bemutatására, rögzítésére képes. Az

"End the Cage Age" elnevezésű európai polgári kezdeményezést az Európai Bizottság elfogadta, ami azt jelenti, hogy a haszonállatok ketreces tartásának fokozatos kivezetésére fog sor kerülni, legkorábban 2027-től. Ez többek között a tojótyúktartást is jelentős mértékben érinti, amellyel kapcsolatban fontos megjegyezni, hogy jelenleg hazánkban a tojótyúk-állomány közel háromnegyede (felújított vagy feljavított) ketreces tartástechnológiában termel. A ketrec nélküli rendszerekben, nagy állományméretek esetén a tyúkok társas viselkedése felértékelődik, a viselkedés folyamatos monitorozására van szükség. Az első kulcsfontosságú lépés a gépi látástechnológia alkalmazása során a tyúkok azonosítása és a felvételeken látható egyedek számának meghatározása. Ezt követheti a viselkedés, a jólléti állapot, az egészségi állapot és egyéb tényezők automatikus nyomon követése. A videófelvételeket Mosonmagyaróváron, a Széchenyi István Egyetem Uni-Agro-Food Kft. sárga magyar tyúk törzstenyészeteként működő baromfitelepén készítettük. A videofelvételek 49 tyúk tevékenységét rögzítették ellenőrzött környezetben, a hangsúlyt az egyes madarak mozgásának felismerésére és a viselkedés előrejelzésére helyeztük fejlett képfeldolgozási technikák segítségével. A kutatásunkban nyílt, madarakra vonatkozó adatbázisokat és nyílt forráskódú annotációs eszközöket használtunk fel a madarak viselkedésének elemzésére specializált adathalmaz létrehozásához. Amellett, hogy nagy mennyiségű nyers felvétel keletkezett, munkánk eredményeképpen kifejezetten tyúkokra koncentráló annotált adathalmaz készült, amely eredményességét az architektúránk kezdeti verziójának működésével bizonyítunk. Jövőbeni terv a detektálás és előrejelzés eredményességének javítása.

***Kulcsszavak**: tyúk, YOLO, precíziós állattenyésztés, mélytanulás, objektumkövetés, objektumdetektálás, mesterséges intelligencia, gépi látás*

1. Introduction

The European Citizens' Initiative "End the Cage Age" has been embraced by the European Commission, signifying that the confinement of farm animals will be phased out from 2027 at the earliest (European Commission 2021). This will also exert a noteworthy influence on the egg industry, where it is crucial to highlight that currently nearly three-quarters of the laying hen population in Hungary is raised in (enriched or enhanced) cage systems. However, alternative systems also present several welfare concerns. Non-cage systems for laying hens, such as free-range and aviary systems, are increasingly adopted globally due to their potential to improve animal welfare by allowing hens to express natural behaviours. Hens in non-cage systems face an increased risk of parasitic infections such as Coccidia, Trichostrongylidae, Heterakis spp., and Capillaria spp., leading to health issues and decreased productivity (Groves, 2021; Nenadović et al., 2022), while also experiencing poor indoor air quality with elevated levels of ammonia and particulate matter that can harm their respiratory health (Rodenburg et al., 2005, Bist et al., 2023). Feather pecking and cannibalism are also more prevalent in alternative systems, often exacerbated by high dust levels (Kittelsen et al., 2022) and inadequate environmental enrichment (Bonnefous et al., 2022).

Nevertheless, the increased group sizes and enhanced capacity for displaying a wider range of behaviours also play a role in the limitations of cage-free housing systems. An important welfare issue in cage-free setups is the significant problem of severe feather pecking. Another aspect that impacts welfare is the elevated occurrence of fractures experienced during the laying phase (Hartcher and Jones, 2017).

Precision Livestock Farming (PLF) systems can provide a real-time, continuous picture of animal welfare and health, and potentially increase production efficiency (Morrone et al., 2022). According to Rowe et al. (2019) most PLF strategies use image analysis to measure welfare in poultry production (42% out of 264 publications), as camera surveillance systems combined with image processing techniques are relatively inexpensive methods to objectively measure poultry behaviour without entering the barn, which results in behavioural changes in the animals. Computer vision, due to the advancement of information technologies in the poultry sector, enables the surveillance of animals without the requirement for human resources and can display and document an abundance of information (Okinda et al., 2020).

In cage-free settings with substantial flock sizes, the social conduct of the hens is esteemed, and ongoing behaviour monitoring is necessary. A range of studies have explored the use of artificial intelligence for video analysis in poultry farming.

In this research, the hens are observed autonomously in the chicken coops by a camera installed on the wall above the hens. The objective is to identify the location and movement patterns of individual birds, in order to forecast behaviour and identify aggressive individuals. The location and movement pattern identification is conducted using advanced image processing methods based on convolutional neural networks (YOLOv8, ResNet architecture) and sequential serial processing (LSTM chaining). Our benchmark model, the ChickTrack system, executes the task using a comparable structure. Given the absence of a dedicated dataset for the task, our goal was to construct a new dataset by utilizing existing open datasets portraying avians (Animal Kingdom, American Birds dataset - NABirds). We also marked the location and kinematic structure (or pose skeleton) on the dataset to identify motion patterns. The marking process used the open-source, freely accessible "Computer Vision Annotation Tool" (CVAT.AI) framework. Besides producing a substantial amount of unedited footage, our efforts led to a marked dataset specifically concentrating on hens, the efficacy of which is illustrated by the functioning of an initial iteration of our structure. Subsequent objectives involve enhancing detection and forecasting capabilities.

2. Materials and Mehods

The research focuses on the detection and classification of the behaviour of chickens, more specifically, the analysis of Yellow Hungarian chickens. The observation and data collection are realized in the realistic setting of an experimental chicken farm in Mosonmagyaróvár, Hungary. The chicken farm is arranged from 32 holding cabinets, each cabinet containing around 50 chickens (hens mixed with roosters). Currently, only a single cabinet is equipped with an exterior camera, observing the chickens continuously, and generating raw video footage for further analysis. The chicken coop and the installation process are depicted in Figure 1. The setting provides the raw footage for further artificial intelligence and data analysis applications.

Figure 1. *Chicken holding cabinet (camera installation, left) with upper view from camera (right)*

The dataset provides information on the localization and predicted kinematic structure (i.e., skeleton) of chickens. Furthermore, the dataset contains localization information of the following prevalent objects in a poultry setting:

- Hens and roosters.
- Watering bowl for the chickens
- Coops holding hen nests.
- Wall boundaries of the cabins.

The creation of such a dataset is essential, as datasets focusing on poultry – and particularly on chicken - are relatively scarce and usually not open for public use. A particular aim of our research is to provide a dataset for precision agriculture applications.
The video capture of chickens follows a straightforward process depicted in Figure 2. The process starts with observing the chicken flock with a camera. The camera is fixed to the ceiling of the holding cabin and can view the whole cabin and coops. Next, the videos are stored on a storage device that can be placed in a cloud or local storage. In the initial phase, local storage was used, and the camera was directly connected to a personal computer.

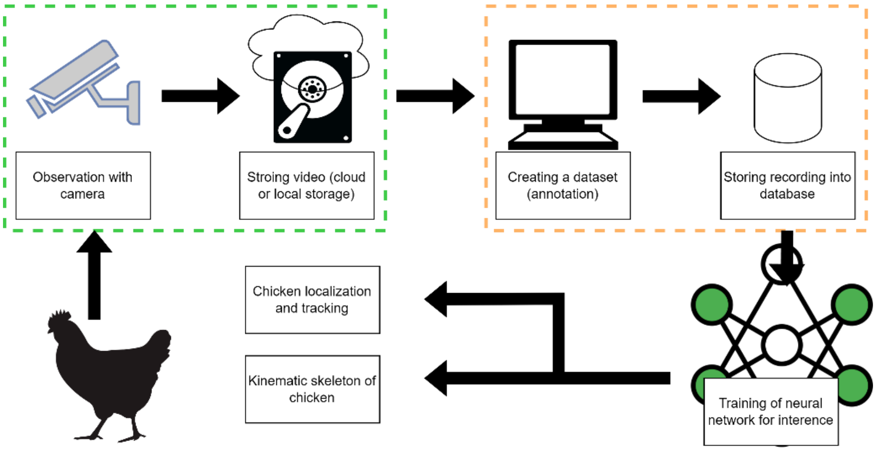

Figure 2. *Closed-loop process of creation of dataset*

The next step is to create the dataset by annotating the captured video, with a tool of selection. After a usable dataset, that can be stored in a database, the neural network utilized for detection can use the dataset for training. The neural network outputs the localization and boundary of chickens on the picture – which can be transformed to their exact 3-dimensional location – as well as their kinematic structure.

2.1. Annotation

The annotation process plays a central role in the current research. Computer Vision Annotation Tool (CVAT.AI) was used as the primary tool for annotation. This framework can be used to create a dataset processable by the neural network (e.g., in a JSON-based format). CVAT.AI enables the production of multiple commonly used outputs (e.g., YOLO, COCO formats) that aid in further training the neural network of choice, especially in the initial development phase.

Besides the typical manual annotation of datasets, CVAT.AI allows the automatic annotation of the datasets. This approach to annotation can use pre-trained networks (e.g., YOLO, Mask-RCNN trained on MS COCO). This ability refines the annotation process, possibly introducing the annotated dataset as an intermediate training dataset used by the automatic annotation. It should be highlighted that this process is ultimately supervised, and human actors can overrule the obtained results. Human supervision is particularly common during the initial training phase. The refined process is depicted in Figure 3. Automatic annotation is typically an offline task that does not require high-speed detection but higher accuracy. Therefore, networks that surpass YOLO in these performance indicators should be used (e.g., RetinaNet, MASK-RCNN) at the cost of slower but more reliable annotation.

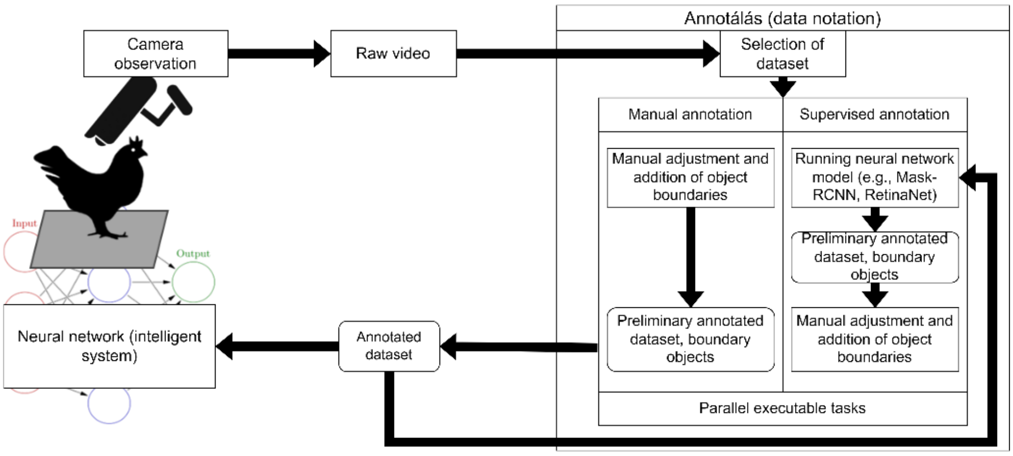

Figure 3. *Refined annotation process with automatic annotation (enabled by tools, e.g. CVAT.AI)*

3. Results and Discussion

This section discusses the main results of our ongoing research: creating an initial dataset resulting from the methodology introduced in Section 2.

3.1. Dataset creation

In the 2.1. subsection, CVAT.AI has been identified as the primary tool of trade for annotation. LabelImg (LabelStudio) was used at the beginning of the work, but this tool seemed cumbersome and lacked features that helped the annotation process. Furthermore, the latest version, as of February 2024, was plagued by frequent crashes and performance problems such as indescribable high memory usage. LabelStudio also struggled to use any video created in any conventional – or even as a raw video – format and codec. For observation, a HIKVision camera was used, outputting videos in H.265 format at a framerate of 25 FPS and with a video resolution of 1080p.

In comparison, CVAT.AI could be used without any reportable performance issue while providing conventional and valuable features such as automatic annotation, clear separation of detectable features, and layered hierarchical structure (handling kinematic pose notation, boundary boxes). Deployment is also based on Docker, providing a straightforward way to system deployment. Automatic annotation could also be used with pre-trained YOLO both on CPU and GPU, enabling possible deployment on a server system. Videos could also be used efficiently and have been treated as image sequences – some videos required proper closing video segments (due to streaming of videos). The computer that ran both annotation frameworks was a personal computer with a Ryzen 5 CPU, 32 GB memory, NVIDIA GTX 2070, and plenty of available hard disk space (around 4 TB) as configuration. The devices, dataset parameters, and configuration are listed in Table 1. The neural network training and the inference have also been partially migrated to the Hungarian Supercomputer, Komondor (KIFÜ) using a GPU supported container. This will eventually improve network training time significantly.

Table 1. Configuration used in the observation

Parameter	Value	Purpose
CPU (Annotation)	Ryzen 5, 16 cores	Annotation
Memory	32 GB	Annotation
Storage	4 TB	Annotation
Annotation framework	CVAT.AI	Annotation
h.265 Camera	HIKVision, 25 FPS	Observation
Raw dataset size	30 GB	Observation
Dataset objects	Hen, chicken, coop, waterbowl, cabin boundary	Observation

With these features, creating the dataset was agile. An average of 500 pictures could be manually annotated daily and two videos automatically. The review of automatic annotation results could be performed with similar performance.

In the current, still rather initial phase, a video quantity of 30 GB has been used as a raw video source. This data amount was enough to produce some preliminary results: bounding boxes and some kinematic pose detection as well (depicted in Figure 4). Naturally, not all of the image frames were usable for annotation, and this made tracking particularly difficult due to the following problems:

- Chickens follow a stochastic and highly unpredictable path, making tracking difficult. Therefore, automatic tracking results had to be overridden.
- Chickens tend to stay in groups, therefore selecting the boundary of individual chickens is difficult, in some cases does not enable usable results.

Figure 4. *Annotation results (left: kinematic key points, right: bounding boxes for YOLO)*

3.2. Architecture design

ChickTrack (Neethirajan, 2022) has been used as the reference architecture. During the research, some problems were identified with this network:

- For the tracking of the chicken location in the picture, Kalman-filtering is used for prediction. As the chicken trajectory is highly unpredictable, this could be a problem. LSTM (Hochreiter and Schmidhuber, 1997) has been successfully used for stochastic sequences (e.g., texts, pose detection, waveform analysis).

- The backbone is not composed of ResNet (He et al., 2015), only PANet, which is an integral part of newer YOLO architectures. YOLO is an architecture consistently developing, with the latest version reaching the 10th iteration (Wang et al., 2024). ResNet could increase the accuracy of detection and tracking as well.
- ChickTrack detects object boundaries and tracks simple vision features, but does not result in the kinematic state of individual chickens.

The current research addresses these problems by extending YOLO architecture with a ResNet backbone and LSTM output for series detection to detect kinematic key points composing the kinematic skeleton (comparison in Figure 5).

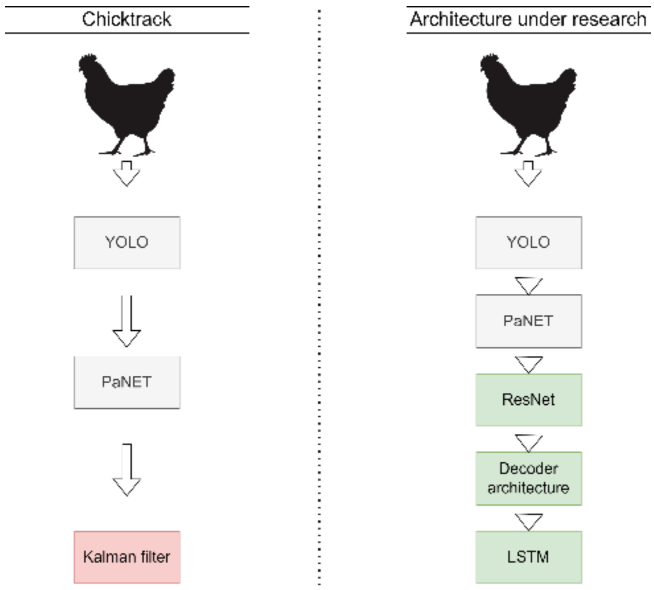

Figure 5. *Comparison of neural architectures useful for tracking chicken behaviour (left: ChickTrack, right: current research)*

4. Conclusions

In our study, the identification and localization of hens within the dataset have been accomplished, and the kinematic key points for motion skeleton recognition have been defined. The system's architecture has been formulated, and the essential infrastructure has been arranged. The dataset is set to expand to encompass kinematic skeleton identification. Presently, the annotation for the kinematic skeleton still needs to be completed. Evaluation is underway on neural networks suitable for kinematic skeleton detection. A tailored architecture is undergoing training and development. Comparisons are drawn with ChickTrack based on measurements. Our forthcoming agenda entails enhancing the dataset and the neural network architecture, followed by conducting tests on various architectures.

Acknowledgements

The authors thank their colleagues at the chicken farm for their assistance with the birds and the camera installation. We acknowledge KIFÜ (Governmental Agency for IT Development, Hungary, https://ror.org/01s0v4q65) for awarding us access to the Komondor HPC facility based in Hungary.

References

Bist, R., Yang, X., Subedi, S., Sharma. M. K., Singh, A., Ritz, C. W. K., Kim, W.K., Chai, L. 2023. Temporal Variations of Air Quality in Cage-Free Experimental Pullet Houses. *Poultry*. **2** (2) 320–333. https://doi.org/10.3390/poultry2020024

Bonnefous, C., Collin, A., Guilloteau, L. A., Guesdon, V., Filliat, C., Réhault-Godbert, S., Leterrier, C. 2022. Welfare issues and potential solutions for laying hens in free range and organic production systems: A review based on literature and interviews. *Frontiers in Veterinary Science*. **9** 952922. http://doi.org/10.3389/fvets.2022.952922

European Comission. 2021. Communication from the Commission on the European Citizens' Initiative (ECI) "End the Cage Age". https://ec.europa.eu/transparency/documents-register/detail?ref=C(2021)4747&lang=en (2024.06.11.)

Groves, P. J. 2021. Non-infectious diseases and laying hen welfare. Animal *Production Science*. **61** 1013–1017. https://doi.org//10.1071/AN19680

Hartcher, K. M., Jones, B. 2017. The welfare of layer hens in cage and cage-free housing systems. World's Poultry Science Journal. 73 (4) 767–782. https://doi.org/10.1017/S0043933917000812

He, K., Zhang, X., Ren, S., Sun, J. 2015. Deep Residual Learning for Image Recognition. 2016 IEEE *Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. https://doi.org/10.1109/cvpr.2016.90

Hochreiter, S., Schmidhuber, J. 1997. Long Short-Term Memory. *Neural Computation.* **9** (8) 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735

Kittelsen, K. E., Tahamtani, F. M., Moe, R. O., Gretarsson, P., Vasdal, G. 2022. Flock Factors Correlated with Elevated Mortality in Non-Beak Trimmed Aviary-Housed Layers. *Animals*. **12** (24) 3577. https://doi.org/10.3390/ani12243577

Morrone, S., Dimauro, C., Gambella, F., Cappai, M.G. 2022. Industry 4.0 and Precision Livestock Farming (PLF): An up to Date Overview across Animal Productions. *Sensors.* **22**, 4319. https://doi.org/10.3390/s22124319

Neethirajan, S. 2022. ChickTrack - A quantitative tracking tool for measuring chicken activity. *Measurement.* **191**, 110819 https://doi.org/10.1016/j.measurement.2022.110819

Nenadović, K., Vučinić, M., Turubatovic, L., Becskei, Zs., Geric, T., Ilić, T. 2022. The effect of different housing systems on the welfare and the parasitological conditions of laying hens. *Journal of the Hellenic Veterinary Medical Society*. **73** (3) 4493–4504. https://doi.org/10.12681/jhvms.27585

Okinda, C. 2020. A review on computer vision systems in monitoring of poultry: A welfare perspective, *Artifical Intelligence in Agriculture*. **4** (1) 184–208. https://doi.org/10.1016/j.aiia.2020.09.002

Rodenburg, T. B., Tuyttens, F. A. M., Sonck, B., De Reu, K., Herman, L., Zoons, J. 2005. Welfare, Health, and Hygiene of Laying Hens Housed in Furnished Cages and in Alternative Housing Systems. *Journal of Applied Animal Welfare Science*. **8** (3) 211–226. https://doi.org/10.1207/s15327604jaws0803_5

Rowe, E., Dawkins, M.S., Gebhardt-Henrich, S.G. 2019. A Systematic Review of Precision Livestock Farming in the Poultry Sector: Is Technology Focussed on Improving Bird Welfare? *Animals*. **9** (9) 614. https://doi.org/10.3390/ani9090614

Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., Ding, G. 2024. YOLOv10: Real-Time End-to-End Object Detection. *ArXiv*. abs/2405.14458. https://doi.org/10.48550/arXiv.2405.14458