Optimized HED- YOLO algorithm for posture detection of pigs

Saraswathi Sivamani[1], Sun Il Chon[1], Ji Hwan Park[1], Seong Ho Choi[2], Dong Hoon Lee[3]*

[1] ThinkforBL Consultancy Services, Seoul, 06236, Rep of Korea;
[2]Department of Animal Science, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, 28644, Rep. of Korea;
[3]Department of Biosystems Engineering, Chungbuk National University, Cheongju-si, Chungcheongbuk-do, 28644, Rep. of Korea;

*Corresponding author

***Abstract*: Precise detection of the object is critical for posture recognition of pigs. The posture detection has been a great interest for researchers, as the detection are linked with the health and welfare of the livestock. Thus, the images obtained from different angles and distance making the posture detection a challenging task. In this paper, we propose an effective posture detection with the combination of (YOLO) and holistically-nested edge detection (HED) to address the issues. The proposed model analysis whether the image along with the Holistically edge detection and machine learning approach could be utilities to identify the posture detection of the pig such as sitting, lying and standing postures. Two approaches that included YOLO with RGB images and Yolov3 with Holistically-Nested Edge Detection (HED) images were included to compare the results. Data from single farm at different time were used for training and validation of the proposed models. The experiment result confirms that the YOLO with HED images was able to detect the posture of the pigs with high accuracy of Mean average precision with 93%.**

***Index Terms*: Image Classification, Animal Behavior, YOLO, Posture Detection, Holistically Edge Detection**

I. INTRODUCTION

Research related to the posture detection of livestock contributes to the assessment of animal behavior. Monitoring the livestock also benefits in time consumption, production management and health of the animal. Implementation of machine learning technique has offered considerable benefits in monitoring the animal behaviour [1-3]. Many researches focus on the automatic detection of animal behaviour image processing, for example Gronskyte et al [4] used image processing group patterning behaviour, while Chen et al [5] analyses the aggressive behaviour using depth sensor and Stavrakakis et al [6] identifies the locomotory behaviour of livestock. Recent studies focus on the behavioural pattern such as lying and standing postures. Nasirahmadi et al [7] involves on behavioural pattern research on pigs which detects both lying and standing posture of the pigs. Olsen et al [8] and Sivamani et al [9] explains the relationship of posture patterns with the health and locomotive changes. Lee et al [10] used Kinect sensor to monitor the standing pigs by removing the noise using spatiotemporal interporal method. Other methods used to detect the posture were ellipse fitting over animal [11], pixel movement [12] and Kinect sensors [13]. The Kinect cameras are even used to detect the weight of the pigs [14-16] and also detect the standing posture of the pig [17]. However, machine learning technique plays major part in the automation object detection [18] and there are many researches that includes the deep learning [19-20]. Similarly, there are relevant attempt on machine learning technique to obtain accurate detection [21-24]. Dynamic filter selection in convolution network assist the invariance at different farm situation [25-26].

Due to the image noises and resolution of the images, the challenges were faced using the machine learning technique. Especially when there are many pigs in an image from different distances and muddy environment. In the fast-growing machine learning field, there are many algorithms that can be resolve the issues. Some of the object detection algorithm used in posture detection are convolution algorithms. Yang et al [27]and Nasirahmadi et al [28] uses the Faster RCNN to detect the pig posture in locomotion and ResNet101 to detect the posture of the pig respectively. Therefore, the principle approach to object detection has always been the CNN algorithms. Although, the result produces high accuracy result, they will inevitably fail to detect under low resolution or noisy situation. In order to solve the issue, we propose the HED masking with the YOLO algorithm for object detection. HED identifies fine and coarse structure of the pig posture, we can use the HED image to train and detection the images accurately.

In this study, we mainly focus on the two type of technique for detecting the posture detection: (1) Due to the complexity of the different breed of pigs in the farm and the unexpected image noises, we focus on the structure of the pigs through the HED images (2) To accommodate large amount of dataset and process the images faster, YOLO algorithm is implemented. As it is impossible to monitor the posture of each pig in the large farm manually, we use the state-of-art machine learning, with edge detection technique to monitor and identify the early problem through posture detection. Therefore, we investigate an automatic posture detection with machine learning technique joint with the edge detection.
Contribution of our studies are,
- use the HED to enhance the edge of the pig posture.
- utilize the YOLO algorithm to train the HED images.
- detecting the posture of pig such as sitting, lying and standing with trained model.
- lastly, we present a comparative study to check the accuracy of the proposed model.

II. MATERIALS AND METHODS

Data Collection

The dataset used in this study was obtained from the commercial farm in South Korea. The data was obtained from 9 pens which was divided as cubicle with 2 pigs in each pen. The images used in the paper were recorded by top-side view camera to mainly capture the dog-sitting posture of the pig. Example of images used for development are used for development of the detection model are illustrated in figure 1.

Figure 1 : Example figures used in this study for posture detection

Overview of the model

Most of the available methods face difficulties when different breed of pigs is mixed in the training dataset due to the image noise, colour, patterns and shape of the pig. We have considered the issue along with touching or multiple pig scenario. Experiments and the deep learning training were conducted on Ubuntu 18.10 LTS, NVIDIA Titan V GPU (12GB Graphic Memory), Xeon Silver 4110 x 2EA, 128GB Memory.

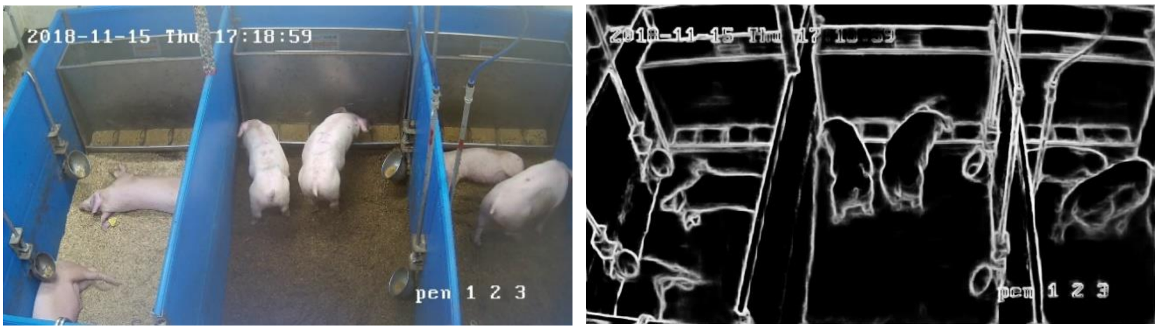
In order to identify the posture such as standing, lying and sitting postures, the images were obtained from farm with all variety of images from different timeline. Total images used for training are 5200 images and 1300 images for validation, which consist of various conditions. From the dataset, 500 images are used for test images different from the training and validation dataset. All the images are converted to HED images with the image size of 1280 x 720. After conversion, all the images are annotated using the annotation tool named LabelImg [29]. The posture such as sitting, lying and standing and incomplete posture such as other pigs are annotated and saved as XML file for each image.

In this study, we used YOLO model on RGB and HED images, aiming to find the best posture detection technique even with the different colour, breed and less clarity. For this, the images were converted to HED image before training with the YOLO algorithm. With HED images, the outline of the object is more vivid for training the images, which increase the accuracy of the object detection.

Holistically-edge Detection

The Holistically-edge Detection is performed with VGG (Visual Geometry Group) network [30]. All the multi-level features are extracted with the VGG's feature extraction capability. The HED is designed similar to the VGG16, which has conv1-2, conv2-2, conv3-3, conv4-3, and conv5-3, with output layers as 5×5, 14×14, 40×40, 92×92 and 196×196, respectively. HED deconvolutes the feature map, resize the image to original image size and the features are fused using convolution network. The main features that helps to get accurate results are perfect location identification and high semantic information from high level network. In only difference between the HED and VGG is remove of fully connected layer in HED network. It helps to minimize the complexity, while providing training to al size of images. While implementing different training method, it improves the cost function which simultaneously train the fusion and output layer. Instead of normal loss function, the edge position loss function is amplified. Thus, optimizing the weights and parameters in the convolution network, with side cost and fusion cost. With these functionality, HED network can be proved to be a good performing network for edge detection, which can be considered as the good target feature information for target detection in this paper. The figure 2 gives a brief illustration of the image in RGB and HED converted, which will be used on this study.

(a)
Figure 2:
RBG image
pen 123 (b)
image of
pen 123



(b)
(a)
of the
HED
the

YOLO3:
YOLO3 [31] has been faster and accurate in detecting every small objects and perfect way to generalize the models. Compared with the previous versions, YOLO is way faster, which can be considered as 1000 times faster than R-CNN and few 100 times faster than Fast R-CNN. Tiny YOLO is also equally faster than the previous version but still lacking in the detection ability. Problem with the small object detection was solved by adding the mutlti-level prediction. With two up sampling features such as 26*16, 52*52 and three detection, the efficiency is increased. In other words, with 52 feature maps, the small targets are easily identified as shown in Figure 3. One more feature that benefits the YOLO was change from softmax loss into logistic loss.

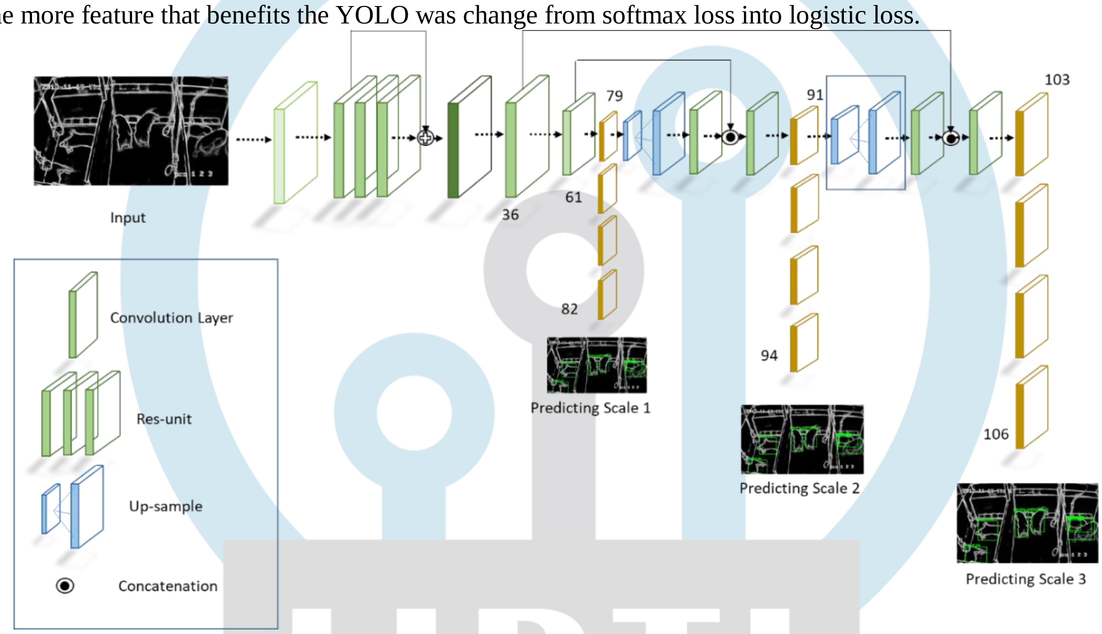


Figure 3: Architecture of the Yolo for object detection.

III. Experimental Results

Monitoring the pig's behavior has always been challenging, mainly in large barns, and labor effective for visual monitoring. The automatic monitoring of the pigs as individual or group has directed to the development of image processing for various benefits. With different condition in farm and environment, the behavior of the pig can change and these precise changes cannot be monitored manually. Therefore, in this study, we perform the machine learning and edge detection technique to identify the induvial pig's behavior with posture detection. Most widely used accuracy calculation average precision and mean average precision were used to calculate the accuracy.

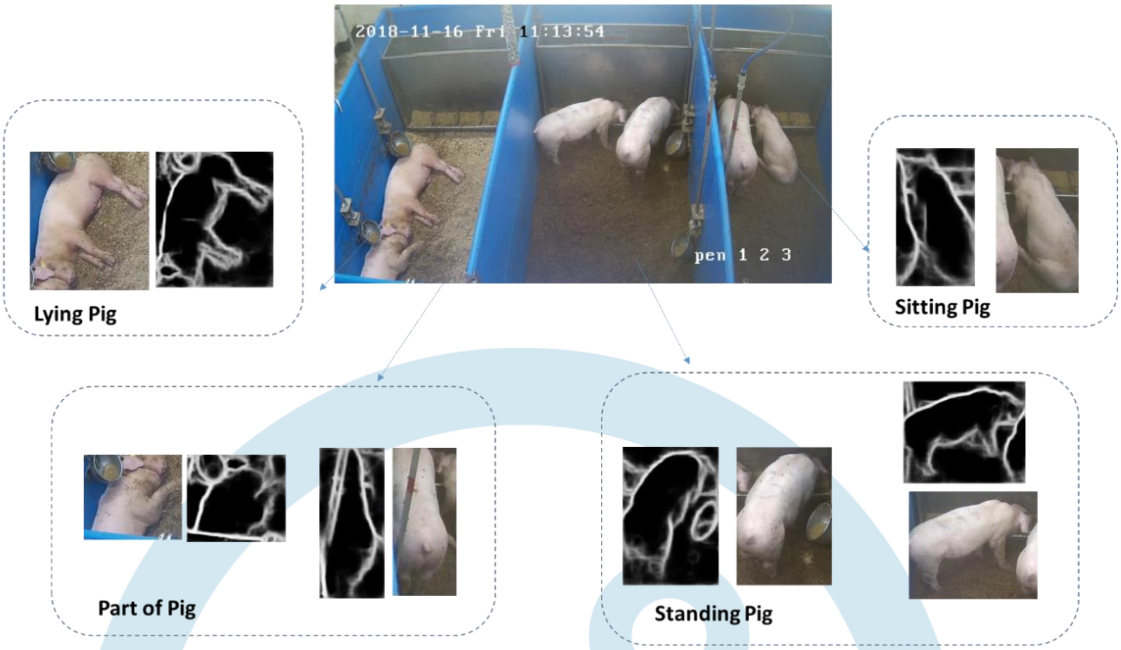


Figure 4: Example image with different posture detection of pig.

Figure 5: Example of detected images of Yolov3 model with RGB image

Figure 5 illustrates examples of detected postures by RGB images with Yolov3 and HED images with Yolov3 models in the different farming conditions in the learning rate of 0.003. As seen from the figure, some of the developed algorithms have the ability to detect different sitting, lying and standing postures under all tested commercial farming conditions. In some cases, there were multiple postures in an image, so the possibility of detecting various postures in an image is also shown in Figure 4. Although the RGB images are gives high accuracy in the posture detection. It may fail to detection the object with different pig colour or blurred images as shown in Figure 5. Most of the pigs were not detected, except the clear objects.

High values for the AP and mAP show the acceptable performance of the YOLO with edge detection approaches for scoring of standing, lying and sitting postures among livestock compared to the other models. In order to get the best result, the learning rate of 0.003 is used for detection models. It can be observed from Table 1 that one of the major misdetections was with respect to standing and sitting postures. This is due to the similarity of posture between the two postures. Therefore, outline of the object is very important to differentiate the posture vividly. HED images provides a vivid outline of the object which is trained to obtain more accurate result.

Nasirahmadi et al. [32] showed that image processing and a linear SVM model was able to score different lying postures (sternal and lateral) in commercial farming conditions. However, the performance of the scoring technique was highly dependent on the output of the image processing method, which led to some wrong scoring in the pigs lying postures. Table 2 shows the comparison of the AP and mAP of the proposed model and the existing methods. HED optimized YOLO model produce 93% accuracy rate compared to the YOLO model with original image. It proves that the HED optimized YOLO detection provides more accuracy compared to other methods by understanding the fine details of the edges in the images. Although the R-FCN shows same rate of accuracy with the proposed model, we have improved the accuracy with the images of unclear background.

Table 1: Confusion Matrix of the proposed HED YOLO detection method

		Predicted Class		
		Sitting	*Standing*	*Lying*
Actual Class	Sitting	1456	416	647
	Standing	552	1789	434
	Lying	645	414	1987

Table 2: Comparison of posture detection methods

Method	AP			mAP
	Sitting	*Standing*	*Lying*	
Yolo (RGB image)	0.87	0.94	0.93	0.90
Yolo (HED image)	0.92	0.93	0.94	0.93
Faster R-CNN [28]	0.86	0.89	0.84	0.86
R-FCN [28]	0.93	0.95	0.92	0.93
SSD [28]	0.76	0.79	0.74	0.76

Using different deep learning approaches, HED with YOLO has proved to have high precision even under the unclear situation. The proposed model can be most effective and valuable model to detect the lying, sitting and standing images, to analysis eh behaviour changes in the livestock. Since the model was created with the data from a single farm, the adaptation is required to identify the posture detection on a wide range.

IV. Conclusion

In this study, we mainly focus on the posture detection of pigs on various circumstance using HED and YOLO technique. From birth to slaughter, monitoring the pig's behaviour is essential to observe the early symptoms of any health issues. However, this can only be realized by state-of-art machine learning technique, which can process the images from different situation. In this study, data from different pen at different period were used to record video data and converted into an image database. The techniques proposed in this study were based on using image data from the surveillance system by using YOLOv3 with both RGB images and HED images. Training and validation images dataset used in the study were 80% and 20% respectively., with a total of around 6000 standing, lying, sitting and standing postures. The trained model was then evaluated using 500 new images. Results of the testing phase showed a high level of mAP and good processing speed for the different postures in the HED optimized YOLO model. The model achieved satisfying result in monitoring the various sitting, lying and standing postures of pigs in image data. Compared to the result with original YOLO model, the HED optimized YOLO model has a high prediction accuracy. The model can be adapted to various similar detection issues, which has proved its flexibility and robustness.

References

[1] A. Nasirahmadi, S. A. Edwards, & B. Sturm. Implementation of machine vision for detecting behaviour of cattle and pigs. Livestock Science, 2017, vol.202, pp. 25-38.

[2] S.G. Matthews, A. L. Miller, J. Clapp, T. Plötz, & L. Kyriazakis, Early detection of health and welfare compromises through automated detection of behavioural changes in pigs. The Veterinary Journal, 2016, vol.217, pp.43-51.

[3] A.R. Frost, R.D.Tillett, S.K.Welch, The development and evaluation of image analysis procedures for guiding a livestock monitoring sensor placement robot. Comput. Electron. Agric. 2000, vol 28, pp.229–242

[4] R. Gronskyte, L. H. Clemmensen, M. S. Hviid, & M. Kulahci. Pig herd monitoring and undesirable tripping and stepping prevention. Computers and Electronics in Agriculture, 2015, vol. 119, pp. 51-60.

[5] C. Chen, W. Zhu, D. Liu, J. Steibel, J. Siegford, K. Wurtz, .. & T. Norton. Detection of aggressive behaviours in pigs using a RealSence depth sensor. Computers and Electronics in Agriculture, 2019, vol.166, pp. 105003.

[6] S. Stavrakakis, W. Li, J. H. Guy, G. Morgan, G. Ushaw, G. R. Johnson, & S. A. Edwards. Validity of the Microsoft Kinect sensor for assessment of normal walking patterns in pigs. Computers and Electronics in AgriSculture, 2015, vol.117, pp. 1-7.

[7] A. Nasirahmadi, S. A. Edwards, S. M. Matheson, & B. Sturm. Using automated image analysis in pig behavioural research: Assessment of the influence of enrichment substrate provision on lying behaviour. 2017, Applied Animal Behaviour Science, vol.196, pp. 30-35.

[8] A. W. Olsen, L. Dybkjær, & H. B. Simonsen. 2000. Behaviour of growing pigs kept in pens with outdoor runs: II. Temperature regulatory behaviour, comfort behaviour and dunging preferences. Livestock production science, 2000, vol.69, no3): 265-278.

[9] S. Sivamani, S. H. Choi, D. H. Lee, J. Park, & S. Chon. 2020. Automatic Posture Detection of pigs on real-time using YOLO framework, 2020, vol.5, no.6, pp. 81-88.

[10] J. Lee, L. Jin, D. Park, & Y. Chung. 2016. Automatic recognition of aggressive behavior in pigs using a kinect depth sensor. Sensors, 16(5): 631.
[11] M. A. Kashiha, C. Bahr, S. Ott, C. P. Moons, T. A. Niewold, F. Tuyttens, & D. Berckmans. 2014. Automatic monitoring of pig locomotion using image analysis. Livestock Science, 2014, vol.159, pp. 141-148.
[12] M. Riekert, A. Klein, F. Adrion, C. Hoffmann, & E. Gallmann. Automatically detecting pig position and posture by 2D camera imaging and deep learning. Computers and Electronics in Agriculture, 2020, vol.174, pp. 105391.
[13] M. Ju, Y. Choi, J. Seo, J. Sa, S. Lee, Y. Chung, & D. Park. 2018. A Kinect-based segmentation of touching-pigs for real-time monitoring. Sensors, 2018, vol.18, no.6, pp. 1746.
[14] I. C. Condotta, T. M. Brown-Brandl, K. O. Silva-Miranda, & J. P Stinn. Evaluation of a depth sensor for mass estimation of growing and finishing pigs. Biosystems Engineering, 2018, vol.173, pp. 11-18.
[15] A. Pezzuolo, M. Guarino, L. Sartori, L. A. González, & F. Marinello. On-barn pig weight estimation based on body measurements by a Kinect v1 depth camera. Computers and Electronics in Agriculture, 2018, vol.148, pp. 29-36.
[16] J. Kongsro. Estimation of pig weight using a Microsoft Kinect prototype imaging system. Computers and Electronics in Agriculture, 2014, vol. 109, pp. 32-35.
[17] S. G. Matthews, A. L. Miller, T. PlÖtz, & I. Kyriazakis. 2017. Automated tracking to measure behavioural changes in pigs for health and welfare monitoring. Scientific reports, 2017, vol.7. no.1, pp. 1-12.
[18] M. L. Everingham, Van Gool, C. K. Williams, J. Winn, & A. Zisserman. 2010. The pascal visual object classes (voc) challenge. International journal of computer vision, 2010, vol.88, no.2, pp. 303-338.
[19] J. V. Huang, Rathod, C. Sun, M. Zhu, A. Korattikara, A. Fathi, ... & K. Murphy. Speed/accuracy trade-offs for modern convolutional object detectors. In Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 7310-7311.
[20] I. Goodfellow, Y. Bengio, A. Courville, & Y. Bengio. Deep learning, 2016, vol. 1, no. 2. Cambridge: MIT press.
[21] E. T. Psota, M. Mittek, L. C. Pérez, T. Schmidt & B. Mote. 2019. Multi-pig part detection and association with a fully-convolutional network. Sensors, 2019, vol. 19, no.4, pp. 852.
[22] N. Kleanthous, A. Hussain, A. Mason, J. Sneddon, A. Shaw, P. Fergus, ... & D. Al-Jumeily. 2018, December. Machine learning techniques for classification of livestock behavior. In International Conference on Neural Information Processing (pp. 304-315). Springer, Cham.
[23] A. Yang, H. Huang, X. Zhu, X. Yang, P. Chen, S. Li, & Y. Xue. 2018. Automatic recognition of sow nursing behaviour using deep learning-based segmentation and spatial and temporal features. Biosystems Engineering, 175: 133-145.
[24] J. Seo, J. Sa, Y. Choi, Y. Chung, D. Park, & H. Kim. 2019, February. A yolo-based separation of touching-pigs for smart pig farm applications. In 2019 21st International Conference on Advanced Communication Technology (ICACT), 2019, pp. 395-401. IEEE.
[25] X. Zhu, C. Chen, B. Zheng, X. Yang, H. Gan, C. Zheng, ... & Y. Xue. 2020. Automatic recognition of lactating sow postures by refined two-stream RGB-D faster R-CNN. Biosystems Engineering, 2020, vol.189, pp. 116-132.
[26] G. Ciaparrone, F. L. Sánchez, S. Tabik, L. Troiano, R. Tagliaferri, & F. Herrera. 2020. Deep learning in video multi-object tracking: A survey. Neurocomputing, 2020, vol. 381, pp. 61-88.
[27] Q. Yang, D. Xiao, & S. Lin. 2018. Feeding behavior recognition for group-housed pigs with the Faster R-CNN. Computers and Electronics in Agriculture, 2018, vol.155, pp. 453-460.
[28] A. Nasirahmadi, B. Sturm, S. Edwards, K. H. Jeppsson,, A. C. Olsson, S. Müller, & O. Hensel. 2019. Deep Learning and Machine Vision Approaches for Posture Detection of Individual Pigs. Sensors, 2019, vol.19, no.17, pp. 3738.
[29] Tzutalin, LabelImg. Git Code. Available online: https://github.com/tzutalin/labelImg, 2015(Accessed on May 2021)
[30] S. Xie, S., & Z. Tu. 2015. Holistically-nested edge detection. In Proceedings of the IEEE international conference on computer vision (pp. 1395-1403).
[31] J. Redmon, & A. Farhadi. Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018.
[32] A. Nasirahmadi, S.A. Edwards, S. M. Matheson, & B. Sturm. 2017. Using automated image analysis in pig behavioural research: Assessment of the influence of enrichment substrate provision on lying behaviour. Applied Animal Behaviour Science, 2017. Vol. 196, pp. 30-35.