Hassan, M., Nordbø, Ø., Olstad, K., Aasmundstad, T. (2026) Automated osteochondrosis detection from CT scanned pigs; genetic parameters and comparison with manual scoring, *Proceedings 13th World Congress on Genetics Applied to Livestock Production* 23969: https://doi.org/10.31274/wcgalp.23969 ©2026 Hassan, et al. CC BY-NC-ND

Automated osteochondrosis detection from CT scanned pigs; genetic parameters and comparison with manual scoring

M.U. Hassan[1,2], Ø. Nordbø[2,3], K. Olstad[1] and T. Aasmundstad[2,3*]

[1]Faculty of Veterinary Medicine, Norwegian University of Life Sciences, Ås, Norway; [2]Norsvin SA, Storhamargata 44, 2317 Hamar, Norway; [3]Topigs Norsvin, Meerendonkweg 25, 5216 TZ 's-Hertogenbosch, The Netherlands; [*]Corresponding author: torunn.aasmundstad@norsvin.no

Introduction

Osteochondrosis (OC) is a developmental orthopedic disease that is defined as a focal disturbance in the process of bone growth by endochondral ossification (Ytrehus *et al.,* 2007). OC lesions can undergo pathologic fracture leading to osteochondrosis dissecans (OCD) fragments that cause joint swelling, pain, leg weakness and lameness/limping in pigs (Ytrehus et al., 2007). At Topigs Norsvin, test boars in the genetic nucleus have been scanned at 120-130 kg using computed tomography (CT) since 2008. In addition to measuring the carcass traits on boar selection candidates, a manual protocol for scoring OC lesions in the elbow and knee joints (Aasmundstad *et al.,* 2013) has been applied. However, scoring lesions is a subjective and labor-intensive task that is vulnerable to human error. Therefore, there is a need for automating the detection and quantification of OC lesions.

Deep learning-based methods like Unet and its variants have proven suited for medical image segmentation tasks (Ronneberger *et al.,* 2015). These methods have been applied in many medical imaging-related problems such as liver segmentation, myocardial infarction, and brain tumor segmentation (Punn and Agarwal, 2022). Since the inception of vision transformers, biomedical image segmentation methods have been significantly improved and widely applied to complex segmentation problems due to their ability to capture long-range dependencies using self-attention mechanisms (Zhou *et al.,* 2023). OC segmentation from 3D CT images is a challenging task. Because the OC lesions are relatively small compared with the CT image size of a 120 kg pig; a full CT image contains about 3×10^8 voxels while on average, a typical OC lesion size is ~300 voxels (Olstad *et al.,* 2022). Additionally, the lesions' sizes and morphology are heterogeneous, which also makes it difficult to capture using traditional image segmentation methods, such as clustering or edge-based methods. However, a specialized self-attention-based transformer model can overcome this challenge. In this study, we apply a transformer-based 3D volumetric segmentation model called nnFormer to segment and quantify the OC lesions in the pig joints (Zhou *et al.,* 2023).

The primary objectives of this study are to develop an automated deep learning-based methodology for detecting and quantifying OC lesions in pig joints from 3D CT images, and to evaluate its effectiveness as an alternative to manual scoring for genetic assessment. The main contributions of this work are as follows: (i) the development and validation of a transformer-based 3D segmentation approach for automated detection and quantification of OC lesions from CT images, and (ii) an evaluation of its feasibility for integration into routine genetic evaluation workflows by estimating genetic parameters in 1,502 Landrace boars and benchmarking performance against the established manual scoring protocol.

Materials & Methods

Animals. 1,502 genotyped Landrace boars born from 2015 and onwards were randomly picked out of our database for the study. The boars were born and raised to about 30 kg in different nucleus herds in Norway, before being sent to the central boar testing station. At the end of test period (at 120-130 kg) the animals are sedated and CT-scanned (see Gjerlaug-Enger *et al.,*

2012 for a more detailed description). Each scan consists of about 1,100 images (dependent of the animals' length) with pixel size equal to 0.9355x0.9355 mm and slice thickness of 1.25 mm. Further, manual OC scorings (Aasmundstad *et al.,* 2013) were extracted from the database and made available for the study.

CT data. Different tissues, including bones and joints in the limbs were segmented and identified, using a combination of deep-learning and heuristic methods (Nordbø *et al.,* 2025). The CT DICOM images were then converted to NIfTI format (https://nifti.nimh.nih.gov/) and cropped into bounding-boxes consisting of 128 pixels in x, y, and z direction, centred around knee, hock, shoulder and elbows for segmenting OC lesions.

Training and testing data. In the training and testing of our deep learning model, we used 198 full CT images dataset from (Olstad *et al.,* 2022) which were cropped into 8 joint-centred sub-images making a sub-dataset of 1,584 cropped 3D images. However, during training we only used that labelled data where at least one lesion was present, shrinking the dataset to 1,181 images in the training and 160 images in the test set.

Segmentation of OC lesions. For the segmentation of OC lesions, we trained nnFormer which is a supervised volumetric segmentation-based 3D transformer (Zhou *et al.,* 2023). nnFormer combines convolutional layers for local feature extraction and transformer blocks for capturing long-range dependencies within volumetric data. This hybrid architecture is well-suited for OC lesions because both fine-grained detail and global anatomical context are required for precise segmentation. We employed the default image preprocessing as in nnU-Net (Isensee *et al.,* 2021) and trained the model from scratch for 1,000 epochs on one NVIDIA T4 GPU.
The volumes of OC lesions per joint were calculated and summed to create two phenotypes: (i) *elbow_and_knee*, which cover the same joints as the manual score, and (ii) *all_joints*, which includes volumes across elbows, knees, shoulders, and hocks. Both traits are measured in mm^3.

Genetic analysis. Automated and manually scored OC traits were analysed using bivariate animal models (Lynch and Walsh, 1998). The boars have all been genotyped at different Illumina SNP-chips (60K, 80K, 50K and 25K). The genotypes were filtered and imputed to a set of 50,424 SNPS, using Fimpute (Sargolzaei *et al.,* 2014), and the genomic relation matrix using the *calc_grm* (Calus and Vandenplas, 2016), using the 1st approach from VanRaden, (2008). In the animal model for manual score, the herd year season (**HYS**), and the technician (**Tech**) were modelled as fixed effects, the live weight (**W**) at scanning as a fixed regression, while genetic effect of the animal (**g**) and the error (**e**) were modeled as random effects (see Eq. 1).

$$\mathbf{OC}_{ijkl}=\mathbf{HYS}_i+\mathbf{W}_j+\mathbf{Tech}_k+\mathbf{g}_l+\mathbf{e}_{ijkl} \quad (1)$$

For the automated phenotypes, the fixed effect of the technician was omitted.

Results & Discussion

The deep learning model nnFormer was trained on labelled dataset and evaluated using the dice score (calculated on probabilities of validation set) and F1-score. During training we achieved a dice score of 0.544 as shown in Figure 1. In addition, we achieved an F1-score of 0.592 after running an inference on the test dataset of 160 images. This is because we had limited examples of severe lesions in our training set, and the Dice score is sensitive to small boundary disagreements. nnFormer is a data-hungry model, which is why the F1 and Dice scores were sub-optimal. On the test dataset, we got 14,616 true positive voxels, with 7,588 false positives and 12,559 false negatives, demonstrating the feasibility of detecting small lesions within large CT volumes.

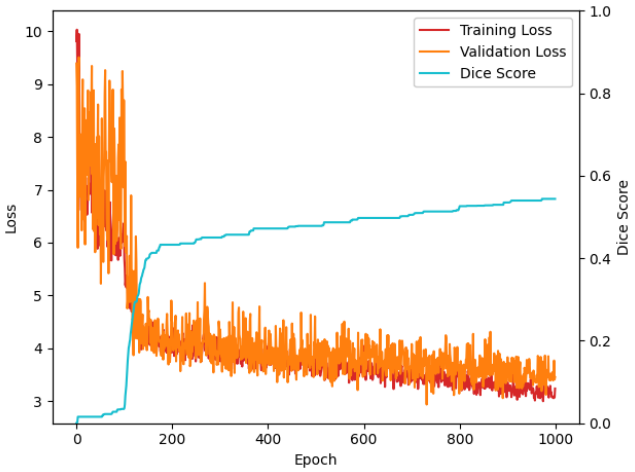

Figure 1. Training performance of nnFormer for 1,000 epochs.

The sum of osteochondrosis lesions was in average 945 mm^3 for *elbow_and_knee* and 1,194 mm^3 for *all_joints*. Both the automated traits were phenotypically correlated with the manual scores as illustrated in Figure 2.

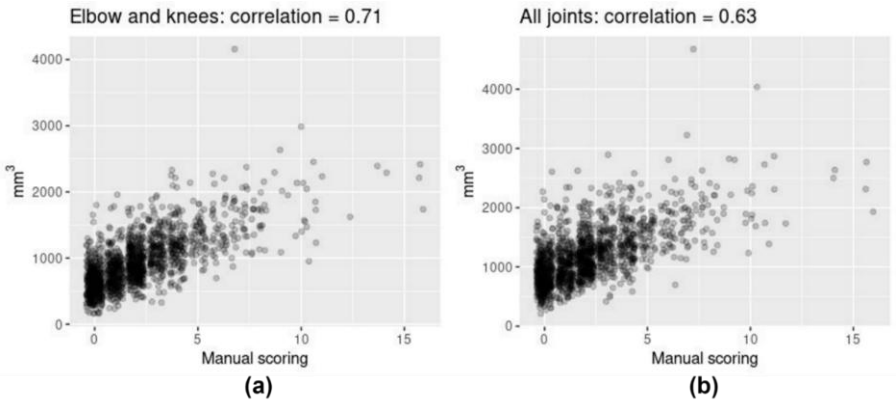

Figure 2. Scatter plots of manual OC scores vs automated traits, (a) elbow_and_knee, and (b) all_joints.

Table 1. Heritabilities (diagonal) and genetic correlations (off-diagonals) between automated and manual scores OC phenotypes.

Elbow and Knee Correlation			All Joints Correlation		
	Elbow and Knee	Manual		All Joints	Manual
Elbow and Knee	0.44 (0.04)		All Joints	0.38 (0.04)	
Manual	0.98 (0.03)	0.25 (0.04)	Manual	0.94 (0.03)	0.26 (0.04)

The automated OC phenotypes had relatively high heritabilities and high genetic correlations with manual scores (Table 1). In particular, the correlation between *elbow_and_knee* scores, as well as manual scores, was high, indicating that these are genetically similar traits. However, the heritability of the automated trait was 76% higher. This is likely due to the higher accuracy in lesion detection and quantification using the automated method, which increases heritability and presents opportunities to improve genetic progress in the breeding program. The *all_joints* trait covers more joints than the other traits analysed. Even though the heritability is somewhat lower than that of the *elbow_and_knee* traits, *all_joints* may capture other aspects of OC and

leg issues that are not captured by traits focusing solely on the *elbow_and_knee*. In the next steps of this research, we will investigate how the different traits correlate with sow longevity, locomotion, and exterior scoring. If the automated method yields higher genetic correlations with sow longevity and other relevant health traits than the manual scores, we can conclude that the automated method provides more efficient, automated, and relevant phenotypes than the manual scoring system.

Conclusions

Automated phenotyping of OC from CT images provides an efficient and accurate method for quantifying OC in the main joints of pigs legs, achieving an F1-score of 0.592. Automated phenotypes had higher heritabilities than those from human scoring: $h^2_{all_joints}$=0.38 (±0.04) and $h^2_{elbow_and_knee}$=0.44 (±0.04) compared with h^2_{manual}=0.26 (±0.04). Genetic correlations between automated and manual scoring were high, at 0.98±0.03 for *elbow_and_knee* and 0.94±0.03 for *all_joints*. These results demonstrate how automated OC phenotypes can contribute to a more efficient breeding program.

Acknowledgements

This study was supported by Research Funds for Agriculture and Food industry (FFL/JA) through project 344491: "Relationship between regional 3D growth and osteochondrosis, and tools for large-scale research into skeletal health, genetics and feeding."

References

Aasmundstad T., Kongsro J., Wetten M., et al. (2013) Animal, 7(10), 1576-1582.
Calus M.P.L., and Vandenplas J. (2016) Calc_grm—a program to compute pedigree, genomic, and combined relationship matrices. ABGC, Wageningen UR Livestock Research.
Gjerlaug-Enger E., Kongsro J., Ødegård J., et al. (2012) Animal, 6(1), 9-18.
Isensee F., Jaeger P.F., Kohl S.A.A., et al. (2021) Nature methods, 18(2), 203-211.
Lynch M., and Walsh B. (1998) Genetics and Analysis of Quantitative Traits. Oxford University Press.
Nordbø Ø., Hassan M.U., Gangsei L.E., et al. (2025) Journal of Animal Science, skaf449.
Olstad, K., Gangsei, L. E., and Kongsro, J. (2022) BMC Veterinary Research, 18(1), 328.
Punn N.S., and Agarwal S. (2022) Artificial Intelligence Review, 55(7), 5845-5889.
Ronneberger O., Fischer P., and Brox T. (2015) U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Cham: Springer international publishing.
Sargolzaei M., Chesnais J.P., and Schenkel F.S. (2014) BMC genomics, 15(1), 478.
VanRaden P.M. (2008) Journal of dairy science, 91(11), 4414-4423.
Ytrehus B., Carlson C.S., and Ekman S. (2007) Veterinary pathology, 44(4), 429-448.
Zhou H.Y., Guo J., Zhang Y., et al. (2023) IEEE transactions on image processing, 32, 4036-4045.