

1 *Type of the Paper (Article)*

2 Happy Cow or Thinking Pig? WUR Wolf – Facial Coding 3 Platform for Measuring Emotions in Farm Animals

4 Suresh Neethirajan^{1*}

5 ¹ Adaptation Physiology Group, Department of Animal Sciences, Wageningen University & Research, 6700
6 AH Wageningen, The Netherlands.

7 * Correspondence: suresh.neethirajan@wur.nl

8 **Abstract:** Emotions play an indicative and informative role in the investigation of farm animal
9 behaviors. Systems that respond and can measure emotions provide a natural user interface in
10 enabling the digitalization of animal welfare platforms. The faces of farm animals can be one of the
11 richest channels for expressing emotions. We present WUR Wolf (Wageningen University & Re-
12 search: Wolf Mascot)—a real-time facial expression recognition platform that can automatically
13 code the emotions of farm animals. Using Python-based algorithms, we detect and track the facial
14 features of cows and pigs, analyze the appearance, ear postures, and eye white regions, and corre-
15 late with the mental/emotional states of the farm animals. The system is trained on dataset of facial
16 features of images of the farm animals collected in over 6 farms and has been optimized to operate
17 with an average accuracy of 85%. From these, we infer the emotional states of animals in real time.
18 The software detects 13 facial actions and 9 emotional states, including whether the animal is ag-
19 gressive, calm, or neutral. A real-time emotion recognition system based on YoloV3, and Faster
20 YoloV4-based facial detection platform and an ensemble Convolutional Neural Networks (RCNN)
21 is presented. Detecting expressions of farm animals simultaneously in real time makes many new
22 interfaces for automated decision-making tools possible for livestock farmers. Emotions sensing
23 offers a vast amount of potential for improving animal welfare and animal-human interactions.

24 **Citation:** Lastname, F.; Lastname, F.;
25 Lastname, F. Title. *Sensors* 2021, 21,
x. <https://doi.org/10.3390/xxxxx>

26 **Keywords:** Biometric verification; facial profile images; animal welfare; precision livestock farm-
ing; welfare monitoring; YOLOv3; YOLOv4; faster RCNN

27 Academic Editor: Firstname
Lastname

28 Received: date
29 Accepted: date
30 Published: date

31 **Publisher's Note:** MDPI stays neu-
32 tral with regard to jurisdictional
33 claims in published maps and insti-
34 tutional affiliations.

1. Introduction

35 Digital technologies, in particular, precision livestock farming, and artificial intelligence
36 have the potential to shape the transformation in animal welfare [1]. To ensure access to
37 sustainable and high-quality health attention and welfare in animal husbandry man-
38 agement, innovative tools are needed. Unlocking the full potential of automated meas-
39 urement of mental and emotional states of farm animals through digitalization such as
40 facial coding systems would help to blur the lines between biological, physical, and dig-
41 ital technologies.



35 **Copyright:** © 2021 by the authors,
36 Submitted for possible open access
37 publication under the terms and
38 conditions of the Creative Commons
39 Attribution (CC BY) license
40 ([http://creativecommons.org/licenses](http://creativecommons.org/licenses/by/4.0/)
41 [/by/4.0/](http://creativecommons.org/licenses/by/4.0/)).

35 Animal caretakers, handlers, and farmworkers typically rely on hands-on observations
36 and measurements while investigating methods to monitor animal welfare. To avoid the
37 increased handling of animals in the process of taking functional or physiological data,
38 and to reduce the subjectivity associated with manual assessments, automated animal
39 behavior and physiology measurement systems can complement the current traditional
40 welfare assessment tools and processes in enhancing the detection of animals in distress
41 or pain in the barn [2]. Automated and continuous monitoring of animal welfare through
digital alerting is rapidly becoming a reality [3].

43 In the human context, facial recognition platforms have long been in use for various ap-
44 plications, such as password systems on smartphones, identification at international
45 border checkpoints, identification of criminals [4, diagnosis of Turner syndrome [5]; de-
46 tection of genetic disorder phenotypes [6]; as a potential diagnostic tool for Parkinson
47 disease [7]; measuring tourist satisfaction through emotional expressions [8]; and quan-
48 tification of customer interest during shopping [9].

49 Emotions

50 Emotions are believed to be a social and survival mechanism that is present in many
51 species. In humans, emotions are understood as deep and complex psychological expe-
52 riences that influence physical reactions. There is an entire sector of science devoted to
53 understanding the sophisticated inner workings of the human brain, yet many questions
54 related to human emotions remain unanswered. Even less scientific research is focused
55 on understanding the emotional capacity of non-human primates and other animals. The
56 ability to interpret the emotional states of an animal is considerably more difficult than
57 understanding the emotional state of a human [10]. The human face is capable of a wide
58 array of expressions that communicate emotion and social intent to other humans. These
59 expressions are so clear that even some non-human species, like dogs, can identify hu-
60 man emotion through facial expression [11]. Each species has its own unique physiolog-
61 ical composition resulting in special forms of expression. Despite human intellectual ca-
62 pacity, emotional understanding of other species through facial observation has proven
63 difficult.

64 Early studies [12] noted the influence of human bias on interpretations and accidental
65 interference with the natural responses of animals. It is not uncommon for humans to
66 anthropomorphize the expressions of animals. The baring of teeth is an example. Hu-
67 mans commonly consider such an expression to be “smiling” and interpret it as a sign of
68 positive emotions. In other species, such as non-human primates, the baring of teeth is
69 more commonly an expression of a negative emotion associated with aggression [13].

70 For these reasons and many others, the involvement of technology is critical in main-
71 taining accurate and unbiased assessments of animal emotions and individual animal
72 identification. In recent years, the number of studies concerning technological interven-
73 tion in the field of animal behavior has increased [10]. The ability of customized software
74 to improve research, animal welfare, the production of food animals, legal identification,
75 and medical practices is astounding.

76 Understanding Animal Emotions

77 The human comprehension of animal emotions may seem trivial; however, it is a mutu-
78 ally beneficial skill. The ability of animals to express complex emotions, such as love and
79 joy, is still being debated within the field of behavioral science. Other emotions, such as
80 fear, stress, and pleasure are studied more commonly. These basic emotions have an
81 impact on how animals feel about their environment and interact with it. It also impacts
82 an animal’s interactions with its counter specifics.

83 Non-domesticated species of animals are commonly observed in the wild and main-
84 tained in captivity to understand and conserve their species. Changes in the natural en-
85 vironment, because of human actions, can be stressful for individuals within a species.
86 Captive non-domesticated animals also experience stress created through artificial envi-
87 ronments and artificial mate selection. If even one animal experiences and displays signs
88 of stress or aggression, its companions are likely to understand and attempt to respond to
89 their emotional state [14]. These responses can result in stress, conflict, and the uneven

90 distribution of resources [15]. The understanding of emotional expression in captive
91 animals can help caretakers determine the most beneficial forms of care and companion
92 matching for each individual, resulting in a better quality of life for the animals in ques-
93 tion.

94 Companion animals are another category of individuals who can benefit from a deeper
95 understanding of animal emotion. Just like humans, individual animals experience dif-
96 ferent thresholds for coping with pain and discomfort. Since many companion animals
97 must undergo voluntary medical procedures for the well-being of their health and their
98 species, it is important to understand their physical responses. Animals cannot tell hu-
99 mans how much pain they are in, so it is up to their caretakers to interpret the pain level
100 an animal is experiencing and treat it appropriately [16]. This task is most accurately
101 completed when the emotions of an animal are clearly and quickly detectable.

102 The understanding of expressions related to stress and pain is impactful in animal agri-
103 culture. Animals used for food production often produce higher quality products when
104 they do not experience unpleasant emotions [17]. The detection of individual animals
105 experiencing stress allows for the early identification of medical complications as well. A
106 study on sows in parturition showed a uniform pattern of facially expressed discomfort
107 during the birthing cycle [18]. In such a case, facial identification of emotional distress
108 could be used to detect abnormally high levels of discomfort and alert human caretakers
109 to the possibility of dystocia.

110 **Facial Recognition Software**

111 Facial recognition software has been used on human subjects for years. It has even con-
112 tributed to the special effect capabilities in films and is used as a password system for
113 locked personal devices. It is a non-invasive method that tracks specific points on an in-
114 dividual's face using photos and videos. These points need not be placed directly on the
115 subject's face; instead, computer software can be customized and trained to identify the
116 location of each point. Once this software identifies an individual's characteristics, it can
117 be modified to detect changes in facial positioning and associate those changes with
118 emotional states.

119 The same method can be used to identify individuals and emotional states when it comes
120 to animal subjects. With a bit of software reconstruction, scientists have been able to cre-
121 ate reliable systems for the assessment of animal emotions through technological means.
122 These systems have been specified to identify multiple species including, cows, cats,
123 sheep, large carnivores, and many species of non-human primate. In studies focused on
124 identifying individual members of the same species within a group, the accuracy of spe-
125 cialized facial recognition software was found to be between 94% and 98.7%. Some of
126 these studies even displayed the ability of software to identify and categorize new indi-
127 viduals within a group and the ability to identify individuals at night [19, 20, 21]. Other
128 studies focused more on the emotional expressions that could be identified through facial
129 recognition software and some of the studies showed an accuracy of around 80% when
130 compared to the findings of professionals in the field of animal emotion identification
131 [22].

132 **The Grimace Scale**

133 The facial recognition software used is based on a series of points in relation to pheno-
134 typic features of the species in question, but it uses an older theory to attach the location
135 of those points to emotional states.

136 The grimace scale is a template created to depict the physical reactions associated with
137 varying levels of discomfort. These scales are created in relation to a specific species and
138 are defined by a numerical scale [23]. In the case of pigs, sheep, and cattle, grimace scales
139 normally focus on tension in the neck, shape of the eye, tension in the brow, nose
140 bunching, and positioning of the ears [24, 18]. These visual cues can be combined with
141 vocal cues to further depict the level of discomfort an animal is experiencing. In species
142 like mice, other expressive physical features must be accounted for, such as whisker
143 movement [25]. For less social species, like cats, the changes in facial expression in re-
144 sponse to pain are more minute but still identifiable with the use of a grimace scale [16].

145 These scales have been proven as an accurate way to assess pain with minimal human
146 bias [23]. They are created through the professional observation of species during con-
147 trolled procedures that are known to trigger pain receptors. Once created, grimace scales
148 can be converted to specific measurements that are detectable through facial recognition
149 software with the assistance of the Viola-Jones algorithm. This algorithm breaks down
150 the facial structure of animal subjects into multiple sections to refine, crop, and identify
151 major facial features [22]. These features make the technological interpretation of animal
152 emotions feasible across a variety of species and in a variety of settings.

153 **Best Way to Manage Animal Emotion Recognition**

154 Studies are most accurate when the spectrum of discomfort, including everything from
155 acute low-grade pain to severe chronic pain, is fully identified. Events of low-grade dis-
156 comfort are significant; however, they may not be identifiable through the production of
157 the stress hormone cortisol [26]. In such situations, discomfort may only be discernable
158 through the facial expressions of an animal detectable by facial recognition software.

159 On large-scale farms, it is important to keep the animals comfortable and relaxed, but it
160 would be impractical and expensive to test the chemical levels of stress present in every
161 animal. The identification of emotional states through facial recognition software pro-
162 vides a more efficient and cost-effective answer. It also provides an opportunity for the
163 identification of very similar individuals in a way that cannot be illegally altered, unlike
164 ear tags, which are sometimes changed for false insurance claims [21].

165 The use of facial recognition software also reduces the need for human interaction with
166 animal subjects. For non-domesticated animals, the presence of human observers can be a
167 stressful experience and alter their natural behavior. Facial recognition software allows
168 researchers to review high-quality video and photo evidence of the subject's emotional
169 expressions without any disturbance. Researchers can even record the identification and
170 actions of multiple individuals within a group of animals at the same time with the help
171 of software like LemurFaceID [20].

172 Room for human error in the form of bias is reduced with the help of facial recognition
173 software. Since humans' experience emotions and have the ability to empathize with
174 other emotional beings, human observers run the risk of interpreting animals' emotional
175 expressions improperly. In a study concerning the pain expressions of sows during par-
176 turation, it was noted that all female observers rated the sows' pain significantly higher
177 than the male observers [18]. That is not to say that one gender is more capable of emo-
178 tional recognition than the other; rather, this situation highlights the opportunity for
179 human error and bias when assessing animals' emotions. With well-calculated software,
180 these discrepancies will cease to exist, and researchers can focus more of their time on
181 finding significant points and connections within recorded data, rather than spending
182 their time recording the data.

183

2. Materials and Methods

184

Dataset

185

Images (Figure 1) and videos of cows and pigs were collected from multiple locations: 3 farms in Canada, 2 farms in the USA, and 1 farm in India. The dataset consisted of 3780 images from a total of 235 pigs and 210 dairy cows. These images were grouped and categorized into multiple subfolders, based on 3 emotions of cows and 6 emotions of pigs. The farm animal's facial expression started from positive to neutral to negative states and returns to neutral state during the data collection process. No elicitation or inducement of affective states on the farm animals were conducted during the data collection.

186

187

188

189

190

191

192

193

Features and Data Processing

194

The collected and grouped images dataset were divided into 9 classes based on the correlation between the facial features such as ear posture and eye whites of cows and pigs and the sensing parameters as compiled in Table 1. The videos and images were pre-processed initially using a 3-stage method (1) Detection of faces, (2) Alignment of faces, (3) Normalization of input. Regular smartphone (Samsung Galaxy S10) was used for capturing images and videos from different angles and directions when the animals were in the barn or pen. The collected data were labelled based on the time stamp and the RFID tags and markers. Faces were not manually extracted but by the MIT LabelImg code [27]. Annotations for labeling different models' bounding boxes were done in the standard format for each: PASCAL format for Faster-RCNN and YOLO format for both YOLOv3 and YOLOv4.

195

196

197

198

199

200

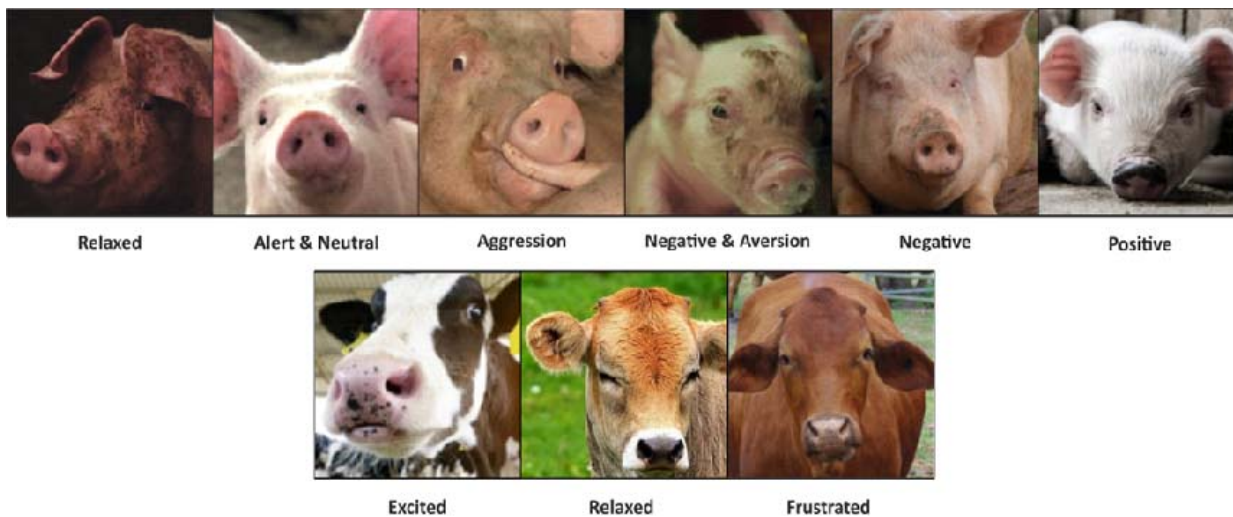
201

202

203

204

205



205

206

207

208

209

210

211

212

213

Figure 1. Sample of images from the data set. Facial features of pigs and cows expressing varying emotions.

214

215

216

Hardware

217

The training and the testing of the 3 models based on YoloV3, YoloV4 and Faster RCNN were performed on NVidia GeForce GTX 1080 Ti graphics processing unit (GPU) running on CUDA 9.0 (compute unified device architecture) and CUDANN 7.6.1 (CUDA deep neural network library), equipped with 3584 CUDA cores and 11 GB memory.

218

219

220

221

222
223
224
225

Table 1. Sensing parameters that were used for each of the 9 classes related to recognizing emotions of cows and pigs [10].

Species Type	Indicators Inferring Emotions	Emotions/Affective States
Cow	Upright ear posture longer	Excited state
Cow	Forward facing ear posture	Frustration (negative emotion)
Cow	Half-closed eyes and ears backwards or hung-down	Relaxed state
Cow	Eye white clearly visible and ears directed forward	Excited state
Cow	Visible eye white	Stress (negative emotion)
Pigs	High frequency ear movement	Stress (negative emotion)
Pigs	Ears forward	Alert (neutral emotion)
Pigs	Ears backward	Negative emotion
Pigs	Hanging ears flipping in the direction of eyes	Neutral emotion
Pigs	Standing upright ears	Normal (neutral state)
Pigs	Ears forward oriented	Aggression (negative emotion)
Pigs	Ears backward and less open eyes	Retreat from aggression or transition to neutral state

226

YOLOv3

227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245

You Only Look Once (YOLO) is one of the fastest Object Detection Systems with a 30 FPS image processing capability and a 57.9% mAP (mean Average Precision) score [28]. YOLO is based on a single Convolutional Neural Network (CNN), i.e. one-step detection and classification. The CNN divides an image into blocks and then it predicts the bounding boxes and probabilities for each block. It was built on a custom Darknet architecture: darknet-19, a 19-layer network supplemented with 11 object detection layers. This architecture, however, struggled with small object detections. YOLOv3 uses a variant of Darknet, a 53-layer Imagenet-trained network combined with 53 more layers for detection and 61.5M parameters. Detection is done at 3 receptive fields: 85×85 , 181×181 , 365×365 , addressing the small object detection issue. The loss function doesn't utilize exhaustive candidate regions but generates the bounding box coordinates and confidence using regression. This gives faster and more accurate detection. It consists of 4 parts, each given equal weightage: regression loss, confidence loss, classification loss, and loss for the absence of any object. When applied to face detection, multiple pyramid pooling layers capture high-level semantic features, and the loss function is altered. Regression loss and confidence loss are given a higher weight. These alterations produce accurate bounding boxes and efficient feature extraction. YOLOv3 provides detection at an excellent speed. However, it suffers from some shortcomings: expressions are affected by the external environment, and orientations/posture are not taken into account.

246

YOLOv4

247 YOLOv4 introduces several features that improve the learning of Convolution Neural
248 Networks (CNNs) [29]. These include Weighted Residual Connections (WRC),
249 Cross-Stage-Partial connections (CSP), Cross mini-Batch Normalization (CmBN), and
250 Self-adversarial training (SAT). CSPDarknet is used as an architecture. It contains 29
251 convolutional layers 3×3 , a 725×725 receptive field, and 27.6M parameters. Spatial
252 Pyramid Pooling (SPP) is added on the top of this layer. YOLOv4 improves the Accuracy
253 Precision Score and FPS of v3 by 10 to 12%. It is faster, more accurate, and can be used on
254 a conventional GPU with 8 to 16 GB-VRAM which enables widespread adoption. New
255 features suppress the weakness and improve on the already impressive face detection
256 capabilities of its predecessor.

257 **Faster R-CNN**

258 Faster R-CNN is the third iteration of the R-CNN architecture. Rich feature hierarchies
259 for accurate detection of objects and features, and semantic segmentation CNN (R-CNN)
260 started in 2014, introducing a method of Selective Search to detect regions of interest in
261 an image and a CNN to classify and adjust them [30]. However, it struggled with pro-
262 ducing real-time results. The next step in its evolution was Fast R-CNN, published in
263 early 2015: a faster model with shared computation capabilities owing to the Region of
264 Interest Pooling technique. Finally came Faster R-CNN, the first fully differentiable
265 model. The architecture consists of a pre-trained CNN (ImageNet) up until an interme-
266 diate layer, which gives a convolutional map. This is used as a feature extractor and
267 provided as input to Region Proposal Network, which tries to find bounding boxes in the
268 image. Region of Interest (RoI) Pooling then extracts features that correspond to the rel-
269 evant objects into a new tensor. Finally, the R-CNN module classifies the contents in the
270 bounding box and adjusts its coordinates to better fit the detected object. Maximum
271 pooling is used to reduce the dimensions of extracted features. A Softmax layer and a
272 regression layer were used to classify facial expressions. This results in Faster R-CNN
273 achieving higher precision and lower miss-rate. However, it is prone to overfitting: the
274 model can stop generalizing at any point and start learning noise.

275 **3. Results**

276 *3.1. Model Parameters*

277 YOLOv3 and YOLOv4 were given image inputs in batches of 64. Learning rate, Mo-
278 mentum, and Step Size were set to 0.001, 0.9, and 20000 steps, respectively. Training took
279 10+ hours for the former and 8+ for the latter. Faster R-CNN accepted input in batches of
280 32. Learning rate, Gamma, Momentum, and Step Size were set to 0.002, 0.1, 0.9, and
281 15000, respectively. It is the most time-consuming to train of the 3, taking 14+ hours. The
282 confusion matrix of DarkNet-53, CSPDarkNet-53, VGG-16 trained and tested on the farm
283 animals' images and videos dataset using YoloV3, YoloV4 and Faster RCNN
284 respectively are shown in Tables S1, S2 and S3.

285 *3.2. Computation Resources*

286 YOLOv3 with its Darknet53 architecture takes the most inference time (0.0331s) com-
287 pared to YOLOv4(0.27s) and Faster R-CNN (0.3s), both of which have CSPDarknet53 and
288 VGG-16 architectures, respectively. YOLOv4 is the computationally efficient model, us-
289 ing 3479 MBs compared to 4759 MBs usage by YOLOv3 and 5877 MBs by Faster R-CNN.
290 YOLOv4 trumps its two competitors when it comes to resources and efficiency, with op-
291 timal memory usage and good-enough inference time. Figure 2 shows the images of farm
292 animals detected by WUR Wolf facial coding platform from the dataset using Faster
293 RCNN technique.

294

295

296

297

298

299

300

301

302

303

304

305

306

307

308

309

310

311

312

313

314

315

316

317

318

319

320

321

322

323



Figure 2. Example emotion detection results from a neutral emotional state pig, and an excited emotional state cow as determined by the WUR Wolf Facial Coding Platform.

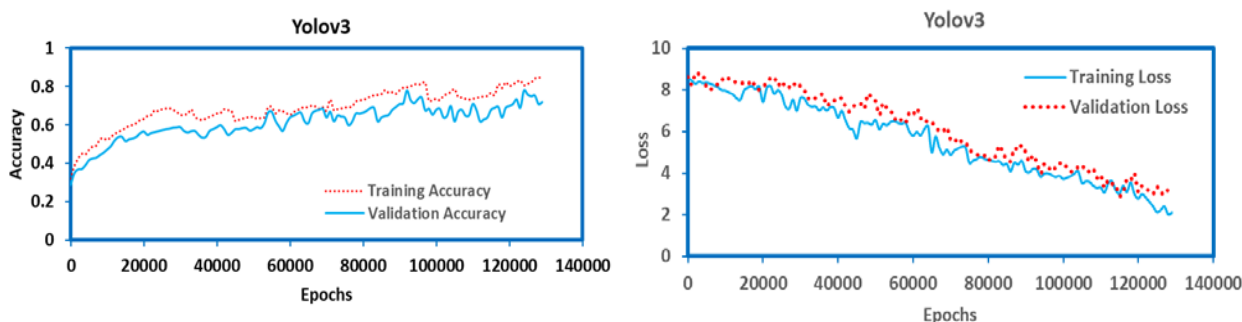


Figure 3. Accuracy and Error Loss for YoloV3

YOLOv3 takes the least amount of time in learning most of the features than the other 2 models and the accuracy curve (Figure 3) flattens earlier as a result. Its fluctuating loss curve is a result of more repetitive predictions and slower convergence as compared to YOLOv4 and Faster R-CNN.

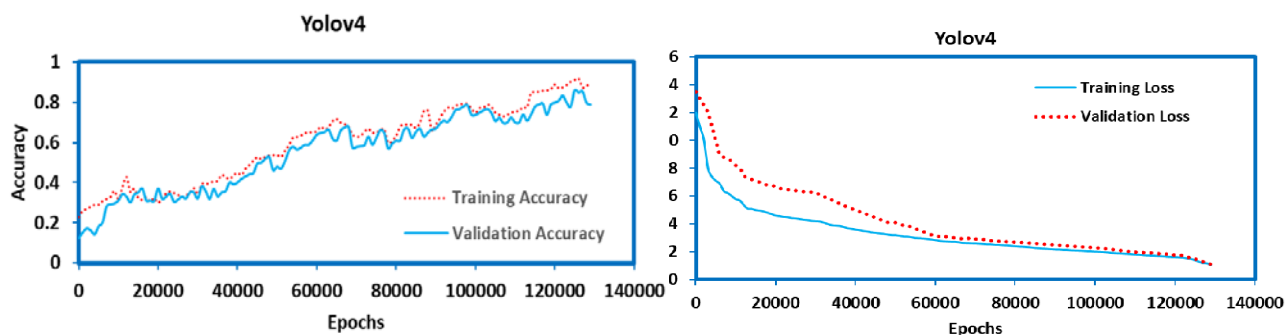


Figure 4. Accuracy and Error Loss for YoloV4

YOLOv4 is slower in learning than Yolov3 but achieves a higher accuracy score and a smoother loss function. Validation accuracy is very close to train accuracy as well, indicating that the model is generalizing well on unseen data (Figure 4) and would perform better in real-time than v3.

Faster R-CNN achieves a higher accuracy (Figure 5) score than both of the YOLO variants as well as converging quickly. However, it performs poorly in generalizing the learning as the difference between validation and train accuracy is very large at multiple times. Faster R-CNN's accuracy score (93.11% on training and 89.19% on validation set) outperforms both YOLOv4 (89.96% on training and 86.45% on validation set) and YOLOv3 (85.21% on training and 82.33% on validation set) on these metrics. Its loss curve is also faster to converge, followed closely by v4, and v3 is the worst performer on this metric.

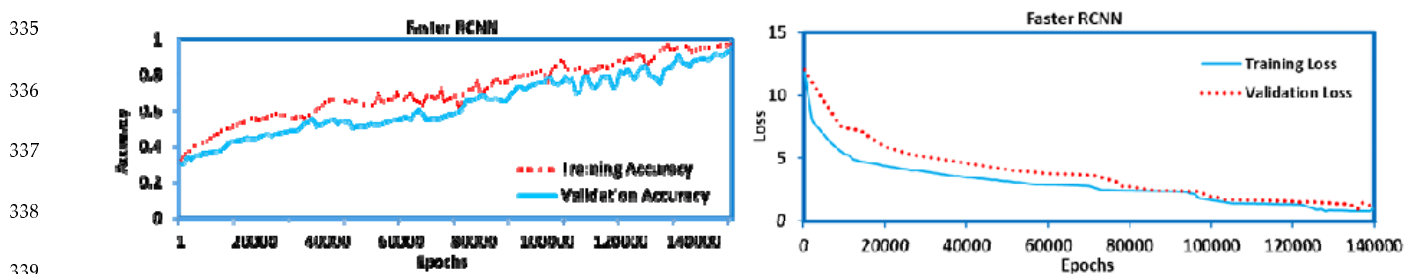


Figure 5. Accuracy and Error Loss for Faster RCNN

Mean Average Precision (mAP)

The mAP score compares the actual bounding box to the detected box and returns a score. The higher the score, the more accurate is the model's object boundary detection. YOLOv4 has a mAP score of 81.6% at 15 FPS, performing better than both the other models. YOLOv3 also performs well on this metric with a mAP score of 77.60% at 11 FPS. Faster R-CNN also provides a moderate mAP score of 75.22%; however, its processing speed is very slow at just 5 FPS. Among the 3, YOLOv4 provides the best bounding boxes at a higher speed.

F1 Score

The annotated labels for both cows and pigs can be grouped on the basis of mental states such as positive, negative, and neutral. Analyzing model performance on these groups is useful in measuring how the model works in different contexts. F1 score is a good measure for this analysis. A Confusion Matrix tabulates the performance of a model on the dataset for which true values are known. Model results are compared against pre-set annotations, and an analysis of them reveals the performance of each model in detecting the emotion portrayed in the picture. Confusion Matrices of all three models are given in the Supplementary Reading Section alongside respective F1 scores. Negative context requires additional effort and reactions, and as a result, there are more pixels with useful information in classification. All 3 models perform return higher True Positives for such cases (Table S4, S5, S6). The average F1 scores of each of the models are as follows: 85.44% for YOLOv3, 88.33% for YOLOv4, and 86.66% for Faster R-CNN. YOLOv4 outperforms the other two in predicting emotion states for each image.

4. Discussion

364 Non-invasive technology that can assess good and poor welfare of farm animals, in-
365 cluding positive and negative emotional states, is possible soon using the proposed WUR
366 Wolf Facial Coding Platform. The ability to track and analyze how animals feel will be a
367 breakthrough in establishing animal welfare auditing tools.

368 In this project, we evaluated the applicability of 3 deep learning-based models for de-
369 termining the emotions of farm animals, Faster R-CNN, and two variants of YOLO:
370 YOLOv3 and YOLOv4. For training the YOLO3 and YOLO4 algorithms, we used the
371 darknet framework. YOLO4 has the CSPDarknet53, while YOLOv3 has the Darknet53.
372 Because of the differences between the backbones, YOLOv4 is faster and provides more
373 accurate results for real-time applications.

374 Demonstration and results of emotion detection of cows and pigs using Faster RCNN
375 (Figure 2) is shown in the attached supplementary video [V1]. Faster RCNN is suitable
376 for mobile terminals where there is a lack of hardware resources in facial expressions
377 recognition [31]. **If speed (time for data processing) is the deciding factor, then YOLOv4 is**
378 **a better choice than Faster RCNN. Due to the advantage of the network design, large**
379 **variations in the dataset composed of facial images and videos with complex and**
380 **multiscale objects is better analyzed by the two-stage Faster RCNN method. Hence, for**
381 **higher accuracy in the results of emotion detection, Faster RCNN is recommended over**
382 **YOLOv4. In on-farm conditions where there may be a lack of equipment related to strong**
383 **data processing ability, Faster RCNN would be a good choice.** Technological advances in
384 the field of animal behavior are a huge step in improving humans' understanding of the
385 animals they share this world with, but there is still room to grow.

386 No facial recognition software created for animals is 100% accurate yet, and so far, only a
387 few common species and non-human primates have had this software modified to iden-
388 tify their physical features. Animal species that are not identified as mammals are mini-
389 mally expressive and have not been tested with facial recognition software for the study
390 of their emotions. One study even brought up the consideration that animals may be able
391 to suppress emotional expression, much like people do in situations where it is socially
392 appropriate to express only certain emotions [25]. There are many questions related to
393 animal emotional expression that have yet to be answered, but there is a good chance that
394 the advancement and implementation of facial recognition software will lead scientists to
395 those answers in the future.

396 5. Conclusions

397 The detailed analysis of the performance of the 3-machine learning python-based models
398 shows the utility of each model in specific farm conditions and how they compare against
399 each other. YOLOv3 learns quickly but gives random predictions and fluctuating losses.
400 Its next iteration, YOLOv4, has improved considerably in multiple regards. If the aim is
401 to balance higher accuracy with faster response and less training time, YOLOv4 works
402 best. If the speed of training and memory usage isn't a concern, the 2-staged Faster
403 R-CNN method performs well and has a robust design for predicting different contexts.
404 The output is accurate, and overfitting is avoided. There is no one-size-fits-all model, but
405 with careful consideration, the most efficient and cost-effective methods can be selected
406 and implemented in automating the facial coding platform for determining farm animal
407 emotions. Facial features and expressions of farm animals provides only
408 one-dimensional aspect of the affective states. Due to the advent of Artificial Intelligence
409 and sensor technologies, in the near future multi-dimensional models of mental and
410 emotional affective states will emerge in the form of measuring behavioural patterns,

411 combined track changes in farm animal postures and behavioural changes with
412 large-scale neural recordings.

413 **Supplementary Materials:** The following are available online at www.mdpi.com/xxx/s1, Figure S1:
414 title, Table S1: Statistical Analysis Results of the Images, Video S1: Demonstration of the WUR Wolf
415 Facial Coding Platform.

416 **Author Contributions:** Conceptualization, S.N.; methodology, S.N.; algorithm development &
417 coding, S.N.; validation, S.N.; formal analysis, S.N.; investigation, S.N.; writing—original draft
418 preparation, S.N.; writing—review and editing, S.N.; supervision, S.N.; project administration, S.N.;
419 All authors have read and agreed to the published version of the manuscript.

420 **Funding:** This research received no external funding.

421 **Conflicts of Interest:** The authors declare no conflict of interest.

422 References

- 423
- 424 1. Neethirajan, S.; Kemp, B. Digital Livestock Farming. *Sensing and Bio-Sensing Res.* **2021**, *32*, 100408.
425 <https://doi.org/10.1016/j.sbsr.2021.100408>.
 - 426 2. Neethirajan, S. Transforming the adaptation physiology of farm animals through sensors. *Animals* **2020**, *10* (9), 1512.
427 <https://doi.org/10.3390/ani10091512>.
 - 428 3. Do, J.P.; Defensor, E.B.; Ichim, C.V.; Lim, M.A.; Mechanic, J.A.; Rabe, M.D.; Schaevitz, L.R. Automated and Continuous Mon-
429 itoring of Animal Welfare through Digital Alerting. *Comp. Med.* **2020**, *70*(4), 313-327.
430 <https://doi.org/10.30802/AALAS-CM-19-000090>
 - 431 4. Purshouse, J.; Campbell, L. Privacy, crime control and police use of automated facial recognition technology. *Crim. L. Rev.*
432 **2019**, *3*, 188-204. ISSN 0011-135X
 - 433 5. Pan, Z.; Shen, Z.; Zhu, H.; Bao, Y.; Liang, S.; Wang, S.; Xiong, G. Clinical application of an automatic facial recognition system
434 based on deep learning for diagnosis of Turner syndrome. *Endocrine* **2020**, 1-9. <https://doi.org/10.1007/s12020-020-02539-3>
 - 435 6. Hadj-Rabia, S.; Schneider, H.; Navarro, E.; Klein, O.; Kirby, N.; Huttner, K.; Grange, D. K. Automatic recognition of the
436 XLHED phenotype from facial images. *Am. J. Med. Genet. A* **2017**, *173*(9), 2408-2414. <https://doi.org/10.1002/ajmg.a.38343>.
 - 437 7. Jin, B.; Qu, Y.; Zhang, L.; Gao, Z. Diagnosing Parkinson Disease Through Facial Expression Recognition: Video Analysis. *J.*
438 *Medical Internet Res.* **2020**, *22*(7), e18697. <https://doi.org/10.2196/18697>.
 - 439 8. González-Rodríguez, M.R.; Díaz-Fernández, M.C.; Gómez, C.P. Facial-expression recognition: an emergent approach to the
440 measurement of tourist satisfaction through emotions. *Telemat. Inform.* **2020**, *51*, 101404.
441 <https://doi.org/10.1016/j.tele.2020.101404>.
 - 442 9. Yolcu, G.; Oztel, I.; Kazan, S.; Oz, C.; Bunyak, F. Deep learning-based face analysis system for monitoring customer interest. *J.*
443 *Ambient Intell. Humaniz. Comput.* **2020**, *11*(1), 237-248. <https://doi.org/10.1007/s12652-019-01310-5>.
 - 444 10. Neethirajan, S.; Reimert, I.; Kemp, B. Measuring Farm Animal Emotions—Sensor-Based Approaches. *Sensors* **2021**, *21*(2),
445 553. <https://doi.org/10.3390/s21020553>.
 - 446 11. Siniscalchi, M.; D'Ingeo, S.; Quaranta, A. Orienting asymmetries and physiological reactivity in dogs' response to human
447 emotional faces. *Learn. Behav.* **2018**, *46*(4), 574–585. <https://doi.org/10.3758/s13420-018-0325-2>.
 - 448 12. Kujala, M.V.; Somppi, S.; Jokela, M.; Vainio, O.; Parkkonen, L. Human empathy, personality and experience affect the
449 emotion ratings of dog and human facial expressions. *PloS one* **2017**, *12*(1), e0170730.
450 <https://doi.org/10.1371/journal.pone.0170730>.

- 451 13. Zhao H.; Li, J.; Wang, X.; Ruliang, P. Facial expression recognition in golden snub-nosed monkeys. *Curr. Zool.* **2020**,
452 695–697. <https://doi.org/10.1093/cz/zoaa055>.
- 453 14. Howarth, E.R.L.; Kemp, C.; Thatcher, H.R.; Szott, I.D.; Farningham, D.; Witham, C.L.; Bethel, E.J. Developing and Validat-
454 ing Attention Bias Tools for Assessing Trait and State Affect in Animals. *Appl. Anim. Behav. Sci.* **2021**, *234*, 1–46.
455 <https://doi.org/10.1016/j.applanim.2020.105198>.
- 456 15. Crump, A.; Bethel, E.; Earley, R.; Lee, V.E.; Arnott, G. Emotion in Animal Contests. *Royal Soc. Pub.* **2020**, 287(1939),
457 20201715. <https://doi.org/10.1098/rspb.2020.1715>.
- 458 16. Finka, L.R.; Stelio, P.L.; Brondani, J.; Tzimiropoulos, Y.; Mills, D. Geometric morphometrics for the study of facial expressions
459 in non-human animals, using the domestic cat as an exemplar. *Sci. Rep.* **2019**, *9*(1), 9883.
460 <https://doi.org/10.1038/s41598-019-46330-5>.
- 461 17. Mota-Rojas, D.; Olmos-Hernandez, A.; Verduzco-Mendoza, A.; Hernandez, E.; Whittaker, A. The Utility of Grimace Scales
462 for Practical Pain Assessment in Laboratory Animals. *Animals* **2020**, *10*(10), 1838. <https://doi.org/10.3390/ani10101838>.
- 463 18. Navarro, E.; Mainau, E.; Manteca, X. Development of a Facial Expression Scale Using Farrowing as a Model of Pain in
464 Sows. *Animals* **2020**, *10*(11), 2113. <https://doi.org/10.3390/ani10112113>.
- 465 19. Guo, S.; Xu, P.; Miao, Q.; Shao, G.; Li, B. Automatic Identification of Individual Primates with Deep Learning Techniques.
466 *iScience* **2020**, *23*, 101412. <https://doi.org/10.1016/j.isci.2020.101412>.
- 467 20. Crouse, D.; Jacobs, R.; Richardson, Z.; Klum, S.; Tecot, S. LemurFaceID: a face recognition system to facilitate individual
468 identification of lemurs. *BMC Zool.* **2017**, *2*(1), 2. <https://doi.org/10.1186/s40850-016-0011-9>.
- 469 21. Kumar, S.; Singh, S.J.; Singh, R.; Singh, A.K. Deep Learning Framework for Recognition of Cattle Using Muzzle Point
470 Image Pattern. *Animal Biometrics* **2017**, *116*, 1–17. <https://doi.org/10.1016/j.measurement.2017.10.064>.
- 471 22. Blumrosen, G.; Hawellek, D.; Pesaran, B. Towards Automated Recognition of Facial Expressions in Animal Models. IEEE
472 International Conference on Computer Vision Workshops (ICCVW), Venice, Italy, **2017**, 2810-2819.
473 10.1109/ICCVW.2017.332.
- 474 23. Mogil, J.S.; Pang, D.S.J.; Dutra, G.G.S.; Chambers, C. The development and use of facial grimace scales for pain measure-
475 ment in animals. *Neurosci. Biobehav. Rev.* **2020**, *116*, 480–493. <https://doi.org/10.1016/j.neubiorev.2020.07.013>.
- 476 24. Guesgen, M.; Beausoleil, N.J.; Leach, M.; Minot, E.O.; Stafford, K.J. Coding and quantification of a facial expression for pain
477 in lambs. *Behav. Processes* **2016**, *132*, 49–56. <https://doi.org/10.1016/j.beproc.2016.09.010>.
- 478 25. Dolensek, N.; Gehrlach, D.A.; Klein, A.S.; Gogolla, N. Facial expressions of emotion states and their neuronal correlates in
479 mice. *Science* **2020**, *368*(6486), 89-94. DOI: 10.1126/science.aaz9468.
- 480 26. Lansade, L.; Nowak, R.; Lainé, A.; Leterrier, C.; Bertin, A. Facial expression and oxytocin as possible markers of positive
481 emotions in horses. *Sci. Rep.* **2018**, *8*(1), 14680. <https://doi.org/10.1038/s41598-018-32993-z>.
- 482 27. Tzatalin, D. LabelImg, Github, 2015. <https://github.com/tzatalin/labelImg>.
- 483 28. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You only look once: Unified, real-time object detection. In Proceedings of
484 the IEEE conference on computer vision and pattern recognition, Las Vegas Nevada, USA, **2016**, 779-788.
485 <https://doi.org/10.1109/CVPR.2016.91>.
- 486 29. Bochkovskiy, A.; Wang, C.Y.; Liao, H.Y.M. Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint*
487 *arXiv:2004.10934*, 2020.
- 488 30. Girshick, R. Fast R-CNN. In Proceedings of the IEEE international conference on computer vision, Santiago, Chile, 2015,
489 1440-1448. <https://doi.org/10.1109/ICCV.2015.169>.

- 490 31. Bao, J.; Wei, S.; Li, J.; Zhang, W. Optimized faster-RCNN in real-time facial expression classification. *IOP Conf. Ser.: Mater.*
491 *Sci. Eng.* **2020**, *790* (1), 012148. DOI: 10.1088/1757-899X/790/1/012148.