1 Machine learning pipeline extracts biologically significant data automatically from avian monitoring

2 videos

3

4 Alex H. H. Chan[1], Jing Qi Liu[1], Terry Burke[2], William D. Pearse[1], Julia Schroeder[1]

5 1 Department of Life Sciences, Imperial College London, United Kingdom

6 2 Department of Animal and Plant Sciences, The University of Sheffield, United Kingdom

7 Correspondence Author: Alex Chan Hoi Hang (thealex.chan2@gmail.com)

8

9 **Abstract**

10 Measuring parental care behaviour in the wild is central to the study of animal life-history trade-offs,

11 but is often labour and time-intensive. More efficient machine learning-based video processing tools

12 have recently emerged that allow parental nest visit rates to be measured using video cameras and

13 computer processing. Here, we used open-source software to detect movement events from videos

14 taken at the nest box of a wild passerine bird population. We show that visit numbers from our

15 automatic data collection pipeline strongly correlate with human observations and predicts an

16 increase in brood fitness. Using a machine learning-assisted annotation approach on a subset of 18

17 videos, we show that the accuracy largely increased and cut annotation time by an average of 5.5x

18 compared to that of a cohort of undergraduate students. Since our automatic pipeline collected

19 biological-meaningful data that would have taken approximately 800 days by human observers, we

20 encourage more researchers to apply existing open-source tools to assist data collection in animal

21 behaviour studies.

23

24

Introduction

Parental care behaviour is a life-history trait that is commonly studied in a wide range of animals (Royle et al., 2012). Parental care is defined as any behaviour that increases the fitness of offspring (Clutton-Brock, 1991; Royle et al., 2012), but often decreases the survival probability of parents (Trivers, 1972) presenting as a life history trade-off (Stearns, 1992). While there are many forms of parental care (*e.g.*, nest building, predator defence, incubation, and feeding; see Royle et al., 2012), parental care in the form of feeding of the young is traditionally considered as very costly for parents (Owens and Bennett, 1994; Winkler and Wilkinson, 1988). As such, a large body of literature focuses on how nest visit frequency with dependent young is associated with aspects of an animal's life history, with a focus on birds, where 90% of species engage in parental care (Cockburn, 2006). For example, work on life history trade-off in terms of parent and offspring fitness (Schroeder et al., 2013), parent coordination (see Ihle et al., 2019; Wojczulanis-Jakubas et al., 2018), parent-offspring conflict (Estramil et al., 2013), or ageing (Wilcoxen et al., 2010) all used the frequency of parental visits to nests (or provisioning) as a proxy of parental investment.

However, measuring parental visitation rates in the wild is labour intensive and time-consuming. Data is traditionally collected by direct observations (*e.g.* Dunn and Cockburn, 1996), which can be invasive by disturbing animals in the vicinity of their nest (Rose, 2009). Less invasive methods include video recording (*e.g.,* García-Navas and Sanz, 2010; Nakagawa et al., 2007), radio tracking (see Mitchell et al., 2012), and the use of radio-frequency tags and antennas at the nest (RFID; Mariette et al., 2011; Ringsby et al., 2009). While automatic methods allow visitation rates to be quantified over long periods of time, the technology is prone to error (up to 20%; see Mariette et al., 2011). On the other hand, video analysis is more flexible and allows for other behaviours to be quantified, like nest defence, copulations or feeding load (Lendvai et al., 2015). Maybe most importantly, using video cameras also allows for visual confirmation of individual identity if colour rings are used. However, manual

2

annotation of video data is time-consuming and error-prone (Tuyttens et al., 2014), calling for more effective ways to extract data from videos (Conway et al., 2021; Weinstein, 2018a).

Recent advances in deep learning (Borowiec et al., 2021) and computer vision (Weinstein, 2018a) allow quick and reliable information to be extracted from field data (Valletta et al., 2017). For example, machine learning methods have been successfully applied to solve problems with species identification (see Wäldchen and Mäder, 2018), bird song complexity measurement (see Pearse et al., 2018; Priyadarshani et al., 2018), social behaviour measurement (see Robie et al., 2017), and individual identification (Bogucki et al., 2019; Ferreira et al., 2020; Körschens et al., 2018; Schofield et al., 2019). Since computing resources are much cheaper than human-labour, such approaches have immense potential in reducing the financial and time costs of data collection, evidenced by a recent increase in popularity for ecological applications (Borowiec et al., 2021; Tuia et al., 2022).

Simultaneously, open-source tools were developed to aid data collection from video data. Examples include software on tracking animals and behaviours in captive settings (Harmer and Thomas, 2019; Pennington et al., 2019; Sridhar et al., 2019; Walter and Couzin, 2021), pose estimation (Graving et al., 2019; Lauer et al., 2021; Pereira et al., 2019), or species identification (Falzon et al., 2020; Van Horn et al., 2015). Particularly, Weinstein (2018b) developed an open-source tool named Deep Meerkat which uses convolutional neural networks (CNNs) to capture movement events from wildlife monitoring videos. Despite the name, the software was initially designed for use with a hummingbird population (Marcot et al., 2019; Weinstein, 2018b), but the software has been adapted for use in marine (Sheehan et al., 2020) and insect (Mertens et al., 2021; Pegoraro et al., 2020) systems. To the best of our knowledge, no literature exists that documents the use of the software in avian systems other than the original hummingbird population.

75 Here, we applied Deep Meerkat (Weinstein, 2018b) to automatically process and extract visitation

76 rates from parental provisioning videos collected in the house sparrow (*Passer domesticus*) population

77 on Lundy Island, UK since 2004 (see Nakagawa et al., 2007). We first compared the results to manually-

78 collected data by trained individuals. Next, we compared the outputs from the automatic processing

79 with alternative methods of annotation, including a machine learning-assisted approach and

80 crowdsourced annotation by a cohort of undergraduate students. Lastly, we determined the biological

81 relevance of the automatic data by testing the hypothesis that increased rates of feeding would lead

82 to increased annual reproductive fitness of the parents – a classic life-history theory hypothesis from

83 behavioural ecology (Trivers, 1972).

84

85

86 **Materials and Methods**

87 **Study System**

88 Data were collected from a population of house sparrows (*Passer domesticus*) on Lundy Island (51° 10'

89 N, 4°40' W) located on the Bristol Channel, UK. The population is part of a long-term study and is

90 monitored systematically since 2000, with >99% of individuals marked with a unique combination of

91 coloured metal rings by the British Trust for Ornithology (Cleasby et al., 2011), and unique passive-

92 integrated transponders (Schroeder et al., 2011). Since house sparrows rarely fly over large bodies of

93 water (Magnussen and Jensen, 2017) very little immigration or emigration has taken place in the

94 population (Schroeder et al., 2015). As a result, the population has high recapture rates with no

95 trapping bias (Simons et al. 2015) and reliable life history data for every individual (Schroeder et al.,

96 2015).

97

98 **Parental provisioning videos**

99 The Lundy sparrow population is situated within a 0.2km^2 area around a small village, since that is the

100 only viable habitat on the island (Schroeder et al., 2011). Nest boxes were put up for the sparrows,

101 and every box is checked systematically to detect all breeding attempts throughout the summer

102 breeding season (Cleasby et al., 2011). After eggs were found and the identity of the parents

103 confirmed by their colour ring combination, 90-minute videos were recorded on the 7th and 11th day

104 after egg hatching, with a video camera 2-5m away from the nest box and a field of view of 30cm

105 radius (see Nakagawa et al., 2007a for detailed procedure) to measure parental visitations.

106

107 When annotating the videos, birds were considered as feeding the young when entering and exiting

108 the nest box, and when feeding behaviour could be seen through the nest box hole without entering

109 the nest box. We recorded when birds are perching outside the nest box but as a separate behaviour

110 such that it did not count towards feeding rate estimates. Feeding rates were then calculated from

111 the first visit of either parent until the end of the video, or until 90 minutes had elapsed from the first

visit, whichever came first. We started counting from the first visit, and not the beginning of the video,

to allow time for the birds to adjust to the disturbance possibly caused by the camera setup. The

resulting time during which visits were scored was termed the effective observational time (Nakagawa

et al., 2007a). The total number of visits by both parents was divided by the effective observational

time to obtain the manual feed rate (feeds per hour) as a measure of parental provisioning.

Between 2004 and 2015, videos (N=2112) were manually annotated by postgraduate students and

researchers, which contributed to multiple publications (see Ihle et al., 2019; Nakagawa et al., 2007;

Schroeder et al., 2019, 2016, 2013, 2012). We used these data as a baseline – the 'manual feed rate'

– to be compared against the automatically-collected data ('automatic feed rate') outlined below.

Automatic Video Processing Pipeline

We processed videos collected between 2011 and 2019 (N=2629) through the open-source

programme Deep Meerkat, which uses convolution neural networks (CNNs) to classify moving frames

and identify moving objects in wildlife monitoring videos (Weinstein, 2018b). Since the performance

of the default model was satisfactory from initial tests, we did not further fine-tune the model using

video frames from our study. We processed the output from Deep Meerkat to merge movement

events detected fewer than 40 frames apart and more than 2 frames in length into the same event.

We then further grouped the events into 7-second video clips, to allow effective post-processing

annotation and matching with manual data. We then tallied the number of events for each video, then

divided the tally by the effective observational time (see above) to obtain the automatic feed rate (in

events per hour; Figure 1). Since most of the feeding events on Lundy sparrows were done by parents

entering the nest box, we assume each feeding event broadly corresponds to 2 detected movement

events by the pipeline (entering and exiting the nest box), hence the automatic feed rate was further

divided by 2. Finally, since certain videos produced massively over-inflated feed rate measures due to

filming environment (*e.g* camera shaking, continuous background movement), we removed any

138 videos that has an automatic feed rate of 72.7 events/hour or more (1.31% videos removed), since

139 that is the max feed rate we have ever recorded from previous manual annotation, assuming any rate

140 more extreme will be an anomaly.

141
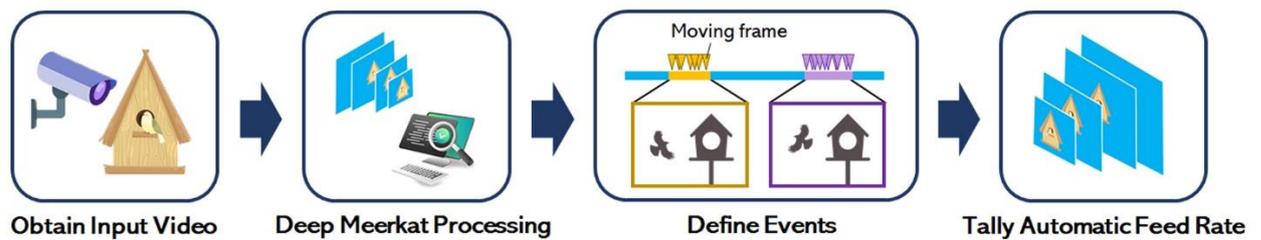
142

143 **Figure 1: Automatic data collection pipeline for processing parental provisioning videos of house**

144 **sparrows on Lundy Island, UK.** Videos were processed using open source software Deep Meerkat

145 (Weinstein, 2018b), outputting frames with movement detected. Events were then defined by

146 grouping movement frames that were less than 40 frames apart and at least 2 frames long. Lastly,

147 events for each provisioning video were tallied to obtain an automatic feed rate

148

149 **Validation**

150 After processing all videos using the automatic pipeline, we first tested the association between the

151 automatic and manual feed rates using Pearson's correlation test. We then chose 18 random videos

152 to be used as a validation dataset, to compare the performance of three different methods of data

153 collection. The first was the automatic feed rate as described above, where videos were processed

154 through the pipeline (Figure 1) without human intervention. The second was a machine learning (ML)-

155 assisted approach, where video outputs of events from the pipeline were further annotated by hand

156 to filter out non-feeding visits. The third was a crowd-sourcing approach, where a cohort of 36

157 undergraduate students were given full-length videos to annotate from scratch. To compare the

158 annotation time between the latter two methods, we also measured the annotation time per video

159 for the ML-assisted and crowd-sourcing approaches.

160

161 To compare the methods quantitatively, we calculated three return rates corresponding to the three

162 methods - the automatic pipeline, machine learning-assisted, and undergraduate return rates. We

163 took the manual feed rate as a baseline value, then calculated return rates by dividing the number of

164 events detected using each method over the manual feed rate for each video (*e.g.* for a given video,

165 if manual feed rate has 10 feeding events and undergraduates found 7, the undergraduate return rate

166 is 0.7). Since a return rate of 1 shows that a method detected the exact number of events as the

167 manual feed rate baseline, we did one sampled t-tests to see whether each return rate was

168 significantly different from this baseline (by setting the theoretical mean (μ) to 1). Lastly, we also

169 compared all methods using pairwise t-tests between the return rates for each annotation method.

170

171 **Case study**

172 Finally, we used the automatic feed rate in a case study to test for the first part of the life-history

173 trade-off predicted by Trivers (1972). We tested whether broods whose parents visited the nest more

174 often had a higher fledgling and recruitment success, using the automatic feed rate. Since each brood

175 was measured on the 7th and 11th day after hatching, we averaged both values to obtain an average

176 feed rate for each brood. However, since feeding rates increases as a brood age (see Schroeder et al.,

177 2019), we also ran the models separately using Day 7 and 11 rates to ensure results are consistent.

178

179 To test the prediction that increased automatic feed rate was associated with higher chick survival in

180 a brood, we fitted generalised linear mixed models with the number of fledglings and the number of

181 recruits for each brood as respective response variables, against automatic feed rates (events per hour)

182 as explanatory variables, using a Poisson link function. A fledgling was defined as a sparrow chick that

183 successfully fledged; recruits were defined as fledglings that produced at least one genetic offspring

184 in their lifetime. To allow effect sizes to be compared, we z-transformed the feed rates. To control for

185 other effects that might affect the fitness metrics, we added the age of the mother and father (see

186   Wiebe, 2018), and hatch date (Days after April 1st) as fixed effects. Since breeding success usually

187   corresponds to peak food abundance (see Cresswell and Mccleery, 2003; Lack, 1968), we also added

188   a quadratic fixed-effect term for hatch date. Next, the population has undergone routine cross-

189   fostering, which is associated with increased survival (see Winney et al., 2015), hence we added a fixed

190   factor for fostered status (yes/no) in all models, however, note that the feed rate we used were always

191   from the parents that did the actual feeding, which are not always the genetic parents of the young.

192   We added the social parent IDs and the year as random effects to control for environmental effects

193   (see Rose et al., 1998) and repeatable feeding rates by individual parents (Nakagawa et al., 2007). The

194   location of the nest box was added as a random effect to control for known environmental effects

195   (Schroeder et al., 2012).

196

197   We ran all models using the R packages MCMCglmm (Hadfield, 2010) on R version 3.6.1 (R Core Team,

198   2013). The posterior distributions and autocorrelations were checked following Hadfield (2014) to

199   ensure all fixed and random effects converged without violating any model assumptions. We defined

200   a parameter estimate as statistically significant if the 95% credible interval did not overlap with zero.

201

202 **Results**

203 **Automatic Video Processing Pipeline**

204 Using the automatic pipeline, we processed a total of 2629 videos and so approximately 3900 hours

205 (160 days) of recordings. We found a significant positive correlation between automatic and manual

206 feed rates (r=0.62, 95%CI: 0.58-0.66, p < 0.001; Figure 2A).

207

208 **Validation**

209 From the 18 validation videos, the automatic pipeline return rate obtained without human

210 intervention was significantly larger than 1 (Table 1), showing that the measure was inflated compared

211 to the baseline set by manual feed rates (Figure 2B). However, the ML-assisted return rate showed

212 that once the pipeline outputs were manually annotated to eliminate non-feeding events, the rate

213 dropped and became not significantly different from 1 and comparable to the baseline (Table 1). The

214 undergraduates performed similarly well, with return rates not significantly different from 1 (Table 1)

215 and not significantly different from the ML-assisted return rate (Figure 2B).

216

217 On average, undergraduate students took 65.4 minutes (min 25 minutes, max 100 minutes) to

218 annotate each video, whereas the ML-assisted approach took an average of 12.0 minutes per video

219 (min 4.6 minutes, max 31.0 minutes), equating to an average of 53.4 minutes saved per video by first

220 processing the videos using the pipeline.
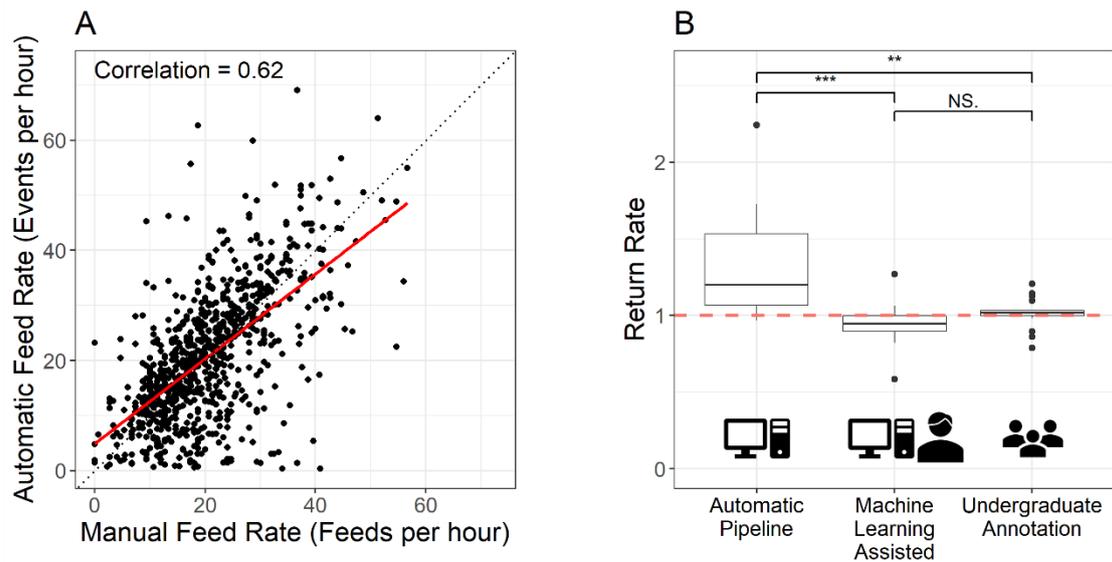
Figure 2: Validation results of automatic video processing pipeline and alternative annotation methods for data collection on house sparrow (*Passer domesticus*) parental provisioning videos on Lundy Island UK. A) Positive correlation between the automatic feed rate collected using the pipeline and manual feed rate collected manually. The red line shows a significant positive correlation, and the dotted line shows a 1:1 line if both rates were equal. B) Comparison of return rates for three separate data collection methods using return rate measures of 18 validation videos. 1) Automatic pipeline: derived from automatic feed rate 2) Machine learning-assisted: further manual annotation of output clips from the pipeline 3) Undergraduate annotation: annotation by a cohort of undergraduate students. The red dotted line represents the reference level of 1, brackets and labels represent the significance of t-test between all groups (N.S: not significant, *: p<0.05, **: p<0.01, ***: p<0.001).

Table 1: Comparison of three annotation methods using a validation dataset of 18 provisioning videos of house sparrows (*Passer domesticus*) on Lundy Island, UK. The return rates were calculated by dividing the number of detected events using each method by the manual feed rate. Test statistics were obtained from a one-sampled t-test, with the theoretical mean (μ) set to 1.

| Metrics	Mean	95% Confidence Interval	t	p-value
A) Automatic pipeline return rate	1.39	1.17-1.63	3.68	0.002
B) Machine learning-assisted return rate	0.94	0.88 – 1.01	-1.78	0.10
C) Undergraduate return rate	1.01	0.97 – 1.06	1.13	0.52

238

239 **Case study**

240 The data that we used contained 658 unique broods, of 2116 individual chicks, of which 728 (34%)

241 fledged and 328 (16%) were recruited. Automatic feed rates significantly predict an increase in the

242 number of fledgelings and recruits for all models fitted (Figure 4; Supporting Information 1). We found

243 a noticeable change in brood survival over the breeding season, with a significant negative quadratic

244 effect for hatch date, showing that fitness is the highest in mid-breeding season (Supporting Table S1).

245
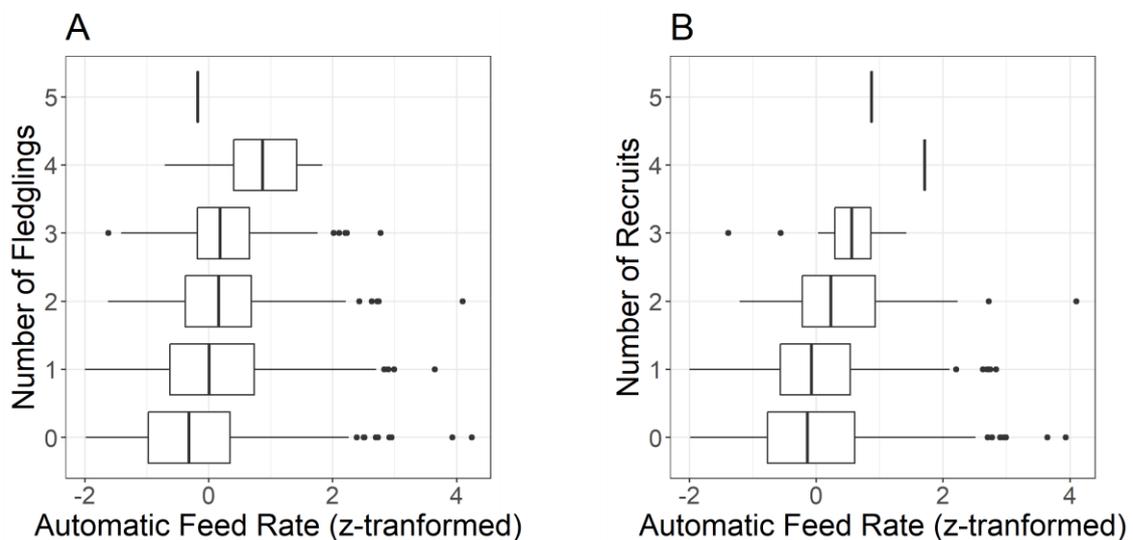
246

247 **Figure 4**: **Increase in the number of (A) fledglings and (B) recruits as average automatic feed rates**

248 **increase in house sparrows (*Passer domesticus*) on Lundy Island, UK** Automatic feed rates were

249 obtained from the pipeline without human intervention. Feed rates were further z_transformed to

250 allow effect sizes to be comparable, with 0 representing the mean and each unit representing one

251 standard deviation away from the mean.

Discussion

Using an automatic data collection pipeline based on Deep Meerkat (Weinstein, 2018b), we extracted visitation rates from parental provisioning videos of house sparrows and found that automatic feed rates positively correlated with manual annotation, and can reproduce biological results, equivalent to ~800 hours (~100 8-hour workdays) of human labour work. Even though the computational time for Deep Meerkat is 1:1 (1 hour video takes ~1 hour to process; see Marcot et al., 2019), computing time is much cheaper than human labour time, especially when techniques like parallel computing were used to further speed up processing. As such, we processed a huge backlog of unprocessed videos from the Lundy sparrow system that would have been infeasible without the use of machine learning methods.

When applying machine learning methods, there is often a trade-off between the variance and bias of a trained model (Geman et al., 1992; but see Neal, 2019). Here, the 18 test videos showed that the pipeline produced biased results, since the obtained automatic feed rate produced overinflated estimates compared to manual feed rates, likely due to over-detection of non-feeding events when birds are still present within the video frame. However, the bias was complemented with low variance (or high consistency), with automatic feed rate correlating well with manual feed rates and predicting an increase in fledgelings and recruits within broods in our case study. Although not comparable with manual feed rates, the automatic feed rate can hence be considered as another proxy of parental investment (see Schroeder et al., 2013; Trivers, 1972) allowing more biological hypotheses to be tested with increased sample size and statistical power. Moreover, the motion detection algorithm of Deep Meerkat is not species-specific, allowing the same method to be applied to other study systems or species, opening a new avenue of research for newly established systems or large video datasets.

In this study, we also tested other methods for data collection, including crowdsourcing undergraduate students and adopting a ML-assisted approach. Manual annotation by undergraduate

278 students was accurate, with mean event return rates slightly higher than the manual feed rates,

279 possibly due to the multiple observer effect (*e.g.* Guay et al., 2013), but can also be due to inflation in

280 detected events due to misidentifications. We also showed that there were no significant differences

281 between the undergraduate and ML-assisted return rates, highlighting the value of using an ML-

282 assisted approach, which was 5.5x faster. Given sex-specific hypotheses on parental investment are

283 often tested in the population (*e.g.* Schroeder et al., 2016), adopting a ML-assisted approach can be

284 useful for collecting additional sex or behavioural data within the feeding clips. Further development

285 using deep learning and computer vision techniques can also classify the sex (see Ferreira et al., 2020)

286 and behaviour (see Conway et al., 2021; Ditria et al., 2021) automatically, which has potential to fully

287 replace human annotation in the future.

288

289 Machine learning approaches are becoming widely used in ecology in recent years (Borowiec et al.,

290 2021). Here, we presented an excellent case study of machine learning assisted data collection that

291 can obtain biologically meaningful results in existing datasets without further manual intervention.

292 With the increase in open-source tools that are being developed in the field to reduce manual

293 annotation efforts (*e.g.* Van Horn et al., 2015; Walter and Couzin, 2021; Weinstein, 2018b), and the

294 increase in computing literacy of ecology graduates (Farrell and Carey, 2018), we encourage

295 researchers to make use of such tools to be adapted to their own existing datasets. This would not

296 only unlock the bottleneck of unanalysed data that would otherwise go to waste, but also allow more

297 interesting hypotheses to be tested.

298

299 **Conclusion**

300 In this study, we used a data collection pipeline based on Deep Meerkat to automatically process

301 parental provisioning videos on Lundy sparrows. Automatic feed rates collected from the machine

302 learning pipeline correlated well with manual annotation, as well as predicting recruitment and

303 fledging success in broods. The pipeline also acts as an effective pre-processing step to allow

304 annotation time to be largely reduced, as we encourage researchers to adopt similar methods in their

305 study systems. Machine learning techniques and open-source tools are becoming widespread in

306 ecology, we provide here a successful case study of the use of such tools to eliminate the bottleneck

307 of laborious manual data collection.

308

309 **Significance Statement**

310 Collecting data on avian parental provisioning rates are often time and labour intensive. Recent

311 advancements in open-source machine learning software allow such data to be effectively collected

312 without human annotation. This paper provides a case-study of the validity and potential of such

313 software to easily collect parental provisioning data in the wild.

314

315 **Conflict of Interest Statement**

316 None

317

318 **Ethics Statement**

319 The study complied with all relevant legislation and guidelines, and had the appropriate ethical

320 approval.

321

331

332   **Author's Contributions**

333   Conception: TB, JS, AC, Funding: TB, JS, Analysis: AC, WP, JS, Writing: AC, TB, JS, WP

334

335

336

References

337 **References**

338 Bogucki, R., Cygan, M., Khan, C.B., Klimek, M., Milczek, J.K., Mucha, M., 2019. Applying deep learning
339   to right whale photo identification. Conserv. Biol. 33, 676–684.
340 Borowiec, M.L., Frandsen, P., Dikow, R., McKeeken, A., Valentini, G., White, A.E., 2021. Deep learning
341   as a tool for ecology and evolution.
342 Cleasby, I.R., Burke, T., Schroeder, J., Nakagawa, S., 2011. Food supplements increase adult tarsus
343   length, but not growth rate, in an island population of house sparrows (Passer domesticus).
344   BMC Res. Notes 4, 431.
345 Clutton-Brock, T.H., 1991. The evolution of parental care. Princeton University Press.
346 Cockburn, A., 2006. Prevalence of different modes of parental care in birds. Proc. R. Soc. B Biol. Sci.
347   273, 1375–1383.
348 Conway, A.M., Durbach, I.N., McInnes, A., Harris, R.N., 2021. Frame-by-frame annotation of video
349   recordings using deep neural networks. Ecosphere 12, e03384.
350 Cresswell, W., Mccleery, R., 2003. How great tits maintain synchronization of their hatch date with
351   food supply in response to long-term variability in temperature. J. Anim. Ecol. 72, 356–366.
352 Ditria, E.M., Jinks, E.L., Connolly, R.M., 2021. Automating the analysis of fish grazing behaviour from
353   videos using image classification and optical flow. Anim. Behav. 177, 31–37.
354 Dunn, P.O., Cockburn, A., 1996. Evolution of male parental care in a bird with almost complete
355   cuckoldry. Evolution 2542–2548.
356 Estramil, N., Eens, M., Müller, W., 2013. Coadaptation of offspring begging and parental
357   provisioning-an evolutionary ecological perspective on avian family life. PloS One 8, e70463.
358 Falzon, G., Lawson, C., Cheung, K.-W., Vernes, K., Ballard, G.A., Fleming, P.J., Glen, A.S., Milne, H.,
359   Mather-Zardain, A., Meek, P.D., 2020. ClassifyMe: a field-scouting software for the
360   identification of wildlife in camera trap images. Animals 10, 58.
361 Farrell, K.J., Carey, C.C., 2018. Power, pitfalls, and potential for integrating computational literacy
362   into undergraduate ecology courses. Ecol. Evol. 8, 7744–7751.
363 Ferreira, A.C., Silva, L.R., Renna, F., Brandl, H.B., Renoult, J.P., Farine, D.R., Covas, R., Doutrelant, C.,
364   2020. Deep learning-based methods for individual recognition in small birds. Methods Ecol.
365   Evol. 11, 1072–1085.
366 García-Navas, V., Sanz, J.J., 2010. Flexibility in the foraging behavior of blue tits in response to short-
367   term manipulations of brood size. Ethology 116, 744–754.
368 Geman, S., Bienenstock, E., Doursat, R., 1992. Neural networks and the bias/variance dilemma.
369   Neural Comput. 4, 1–58.
370 Graving, J.M., Chae, D., Naik, H., Li, L., Koger, B., Costelloe, B.R., Couzin, I.D., 2019. DeepPoseKit, a
371   software toolkit for fast and robust animal pose estimation using deep learning. Elife 8,
372   e47994.
373 Guay, P.-J., McLeod, E.M., Cross, R., Formby, A.J., Maldonado, S.P., Stafford-Bell, R.E., St-James-
374   Turner, Z.N., Robinson, R.W., Mulder, R.A., Weston, M.A., 2013. Observer effects occur
375   when estimating alert but not flight-initiation distances. Wildl. Res. 40, 289–293.
376 Hadfield, J., 2014. MCMCglmm course notes. Available Cran R-Proj.
377   OrgwebpackagesMCMCglmmvignettesCourseNotes Pdf.
378 Hadfield, J.D., 2010. MCMC methods for multi-response generalized linear mixed models: the
379   MCMCglmm R package. J. Stat. Softw. 33, 1–22.
380 Harmer, A.M., Thomas, D.B., 2019. pathtrackr: An r package for video tracking and analysing animal
381   movement. Methods Ecol. Evol. 10, 1196–1202.
382 Ihle, M., Pick, J.L., Winney, I.S., Nakagawa, S., Schroeder, J., Burke, T., 2019. Rearing success does not
383   improve with apparent pair coordination in offspring provisioning. Front. Ecol. Evol. 7, 405.
384 Körschens, M., Barz, B., Denzler, J., 2018. Towards automatic identification of elephants in the wild.
385   ArXiv Prepr. ArXiv181204418.
386 Lack, D.L., 1968. Ecological adaptations for breeding in birds.

387 Lauer, J., Zhou, M., Ye, S., Menegas, W., Nath, T., Rahman, M.M., Santo, V.D., Soberanes, D., Feng,
388   G., Murthy, V., Lauder, G., Dulac, C., Mathis, M., Mathis, A., 2021. Multi-animal pose
389   estimation and tracking with DeepLabCut. bioRxiv.
390 Lendvai, A.Z., Akçay, Ç., Ouyang, J.Q., Dakin, R., Domalik, A.D., St John, P.S., Stanback, M., Moore,
391   I.T., Bonier, F., 2015. Analysis of the optimal duration of behavioral observations based on an
392   automated continuous monitoring system in tree swallows (Tachycineta bicolor): is one hour
393   good enough? PLoS One 10, e0141194.
394 Magnussen, E., Jensen, J.-K., 2017. Ringing recoveries of house sparrow (Passer domesticus) in the
395   Faroe Islands during the years 1963-2007/Ringmerkingar av føroyskum gráspurvum (Passer
396   domesticus) árini 1963-2007. Fróðskaparrit-Faroese Sci. J. 57, 182–189.
397 Marcot, B.G., Lorenz, T.J., Fischer, P., Weinstein, B.G., Cowell, S., 2019. Efficacy of automated
398   detection of motion in wildlife monitoring videos. Wildl. Soc. Bull. 43, 726–736.
399 Mariette, M.M., Pariser, E.C., Gilby, A.J., Magrath, M.J., Pryke, S.R., Griffith, S.C., 2011. Using an
400   electronic monitoring system to link offspring provisioning and foraging behavior of a wild
401   passerine. The Auk 128, 26–35.
402 Mertens, J.E., Brisson, L., Janeček, Š., Klomberg, Y., Maicher, V., Sáfián, S., Delabye, S., Potocký, P.,
403   Kobe, I.N., Pyrcz, T., 2021. Elevational and seasonal patterns of butterflies and hawkmoths in
404   plant-pollinator networks in tropical rainforests of Mount Cameroon. Sci. Rep. 11, 1–12.
405 Mitchell, G.W., Newman, A.E., Wikelski, M., Ryan Norris, D., 2012. Timing of breeding carries over to
406   influence migratory departure in a songbird: an automated radiotracking study. J. Anim.
407   Ecol. 81, 1024–1033.
408 Nakagawa, S., Gillespie, D.O.S., Hatchwell, B.J., Burke, T., 2007. Predictable males and unpredictable
409   females: sex difference in repeatability of parental care in a wild bird population. J. Evol.
410   Biol. 20, 1674–1681.
411 Neal, B., 2019. On the bias-variance tradeoff: Textbooks need an update. ArXiv Prepr.
412   ArXiv191208286.
413 Owens, I.P., Bennett, P.M., 1994. Mortality costs of parental care and sexual dimorphism in birds.
414   Proc. R. Soc. Lond. B Biol. Sci. 257, 1–8.
415 Pearse, W.D., Morales-Castilla, I., James, L.S., Farrell, M., Boivin, F., Davies, T.J., 2018. Global
416   macroevolution and macroecology of passerine song. Evolution 72, 944–960.
417 Pegoraro, L., Hidalgo, O., Leitch, I.J., Pellicer, J., Barlow, S.E., 2020. Automated video monitoring of
418   insect pollinators in the field. Emerg. Top. Life Sci. 4, 87–97.
419 Pennington, Z.T., Dong, Z., Feng, Y., Vetere, L.M., Page-Harley, L., Shuman, T., Cai, D.J., 2019.
420   ezTrack: An open-source video analysis pipeline for the investigation of animal behavior. Sci.
421   Rep. 9, 1–11.
422 Pereira, T.D., Aldarondo, D.E., Willmore, L., Kislin, M., Wang, S.S.-H., Murthy, M., Shaevitz, J.W.,
423   2019. Fast animal pose estimation using deep neural networks. Nat. Methods 16, 117–125.
424 Priyadarshani, N., Marsland, S., Castro, I., 2018. Automated birdsong recognition in complex acoustic
425   environments: a review. J. Avian Biol. 49, jav-01447.
426 R Core Team, 2013. R: A language and environment for statistical computing.
427 Ringsby, T.H., Berge, T., Saether, B.-E., Jensen, H., 2009. Reproductive success and individual
428   variation in feeding frequency of House Sparrows (Passer domesticus). J. Ornithol. 150, 469–
429   481.
430 Robie, A.A., Seagraves, K.M., Egnor, S.R., Branson, K., 2017. Machine vision methods for analyzing
431   social interactions. J. Exp. Biol. 220, 25–34.
432 Rose, A.P., 2009. Temporal and individual variation in offspring provisioning by tree swallows: a new
433   method of automated nest attendance monitoring. Plos One 4, e4111.
434 Rose, K.E., Clutton-Brock, T.H., Guinness, F.E., 1998. Cohort variation in male survival and lifetime
435   breeding success in red deer. J. Anim. Ecol. 979–986.
436 Royle, Nick J, Smiseth, P.T., Kölliker, M., 2012. The evolution of parental care. Oxford University
437   Press.

438   Royle, Nick J., Smiseth, P.T., Mathias, K., 2012. The evolution of parental care, The evolution of
439   parental care. Oxford University Press, Oxford.
440   Schofield, D., Nagrani, A., Zisserman, A., Hayashi, M., Matsuzawa, T., Biro, D., Carvalho, S., 2019.
441   Chimpanzee face recognition from videos in the wild using deep learning. Sci. Adv. 5,
442   eaaw0736.
443   Schroeder, J., Cleasby, I., Dugdale, H.L., Nakagawa, S., Burke, T., 2013. Social and genetic benefits of
444   parental investment suggest sex differences in selection pressures. J. Avian Biol. 44, 133–
445   140.
446   Schroeder, J., Cleasby, I.R., Nakagawa, S., Ockendon, N., Burke, T., 2011. No evidence for adverse
447   effects on fitness of fitting passive integrated transponders (PITs) in wild house sparrows
448   Passer domesticus. J. Avian Biol. 42, 271–275.
449   Schroeder, J., Dugdale, H., Nakagawa, S., Sparks, A., Burke, T., 2019. Social genetic effects (IGE) and
450   genetic intra-and intersexual genetic correlation contribute to the total heritable variance in
451   parental care.
452   Schroeder, J., Hsu, Y.-H., Winney, I., Simons, M., Nakagawa, S., Burke, T., 2016. Predictably
453   philandering females prompt poor paternal provisioning. Am. Nat. 188, 219–230.
454   Schroeder, J., Nakagawa, S., Cleasby, I.R., Burke, T., 2012. Passerine birds breeding under chronic
455   noise experience reduced fitness. PLoS One 7, e39200.
456   Schroeder, J., Nakagawa, S., Rees, M., Mannarelli, M.-E., Burke, T., 2015. Reduced fitness in progeny
457   from old parents in a natural population. Proc. Natl. Acad. Sci. 112, 4021–4025.
458   Sheehan, E.V., Bridger, D., Nancollas, S.J., Pittman, S.J., 2020. PelagiCam: A novel underwater
459   imaging system with computer vision for semi-automated monitoring of mobile marine
460   fauna at offshore structures. Environ. Monit. Assess. 192, 1–13.
461   Sridhar, V.H., Roche, D.G., Gingins, S., 2019. Tracktor: image-based automated tracking of animal
462   movement and behaviour. Methods Ecol. Evol. 10, 815–820.
463   Stearns, S., 1992. The evolution of life histories. Oxford University Press, Oxford.
464   Trivers, R., 1972. Parental investment and sexual selection. Sex. Sel. Descent Man Aldine Gruyter N.
465   Y. 136–179.
466   Tuia, D., Kellenberger, B., Beery, S., Costelloe, B.R., Zuffi, S., Risse, B., Mathis, A., Mathis, M.W., van
467   Langevelde, F., Burghardt, T., 2022. Perspectives in machine learning for wildlife
468   conservation. Nat. Commun. 13, 1–15.
469   Tuyttens, F.A.M., de Graaf, S., Heerkens, J.L., Jacobs, L., Nalon, E., Ott, S., Stadig, L., Van Laer, E.,
470   Ampe, B., 2014. Observer bias in animal behaviour research: can we believe what we score,
471   if we score what we believe? Anim. Behav. 90, 273–280.
472   Valletta, J.J., Torney, C., Kings, M., Thornton, A., Madden, J., 2017. Applications of machine learning
473   in animal behaviour studies. Anim. Behav. 124, 203–220.
474   Van Horn, G., Branson, S., Farrell, R., Haber, S., Barry, J., Ipeirotis, P., Perona, P., Belongie, S., 2015.
475   Building a bird recognition app and large scale dataset with citizen scientists: The fine print
476   in fine-grained dataset collection, in: Proceedings of the IEEE Conference on Computer
477   Vision and Pattern Recognition. pp. 595–604.
478   Wäldchen, J., Mäder, P., 2018. Machine learning for image based species identification. Methods
479   Ecol. Evol. 9, 2216–2225.
480   Walter, T., Couzin, I.D., 2021. TRex, a fast multi-animal tracking system with markerless
481   identification, and 2D estimation of posture and visual fields. Elife 10, e64000.
482   Weinstein, B.G., 2018a. A computer vision for animal ecology. J. Anim. Ecol. 87, 533–545.
483   Weinstein, B.G., 2018b. Scene-specific convolutional neural networks for video-based biodiversity
484   detection. Methods Ecol. Evol. 9, 1435–1441.
485   Wiebe, K.L., 2018. Age-related improvements in fecundity are driven by the male in a bird with
486   partially reversed sex roles in parental care. Oecologia 188, 1095–1104.
487   Wilcoxen, T.E., Boughton, R.K., Schoech, S.J., 2010. Older can be better: physiological costs of
488   paternal investment in the Florida scrub-jay. Behav. Ecol. Sociobiol. 64, 1527–1535.

489 Winkler, D.W., Wilkinson, G.S., 1988. Parental effort in birds and mammals: theory and
490   measurement. Oxf. Surv. Evol. Biol.
491 Winney, I., Nakagawa, S., Hsu, Y.-H., Burke, T., Schroeder, J., 2015. Troubleshooting the potential
492   pitfalls of cross-fostering. Methods Ecol. Evol. 6, 584–592.
493 Wojczulanis-Jakubas, K., Araya-Salas, M., Jakubas, D., 2018. Seabird parents provision their chick in a
494   coordinated manner. PloS One 13, e0189969.
495