1 Machine learning pipeline extracts biologically significant data automatically from wildlife

2 monitoring 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 fledglings and recruits in a brood. We further tested other annotation methods on a subset

17 of videos, showing that a machine learning assisted annotation approach can largely increase the

18 accuracy of the obtained measures and cut annotation time significantly compared to a cohort of

19 undergraduate students. Since our automatic pipeline collected biological-meaningful data that would

20 have taken approximately 800 days by human observers, we encourage more researchers to apply

21 existing open-source tools to assist data collection in animal behaviour studies.

22 **Keywords:** Parental Care, Computer Vision, Deep Meerkat, House Sparrow, Machine Learning,

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 the frequency with which parents visit their nest with dependent young is associated with aspects of an animal's life history, especially in 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

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, the software has not been used 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 in the past. Next, we compared the outputs from the automatic

79   processing with other methods of annotation, including crowdsourced annotation by a cohort of

80   undergraduate students, as well as a machine learning-assisted approach. Lastly, we determined the

81   biological relevance of the extracted data by testing the hypothesis that increased rates of feeding

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

83   hypothesis from 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 colour rings, metal ring by the British Trust for Ornithology (Cleasby et al., 2011), and unique PIT

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

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

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

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

96

97 **Parental provisioning videos**

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

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

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

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

102 confirmed by video, 90-minute videos were recorded on the 7th and 11th day after egg hatching, with

103 a video camera 2-5m away from the nest box and a field of view of 30cm radius (see Nakagawa et al.,

104 2007a for detailed procedure) to measure parental visitations.

105

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

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

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

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

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

111 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 provision rate.

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 post-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 manually-generated data. We then tallied the number of events for each video, then divided the tally by the total time of the video recording to obtain the automatic feed rate (in events per hour; Figure 1). Since the majority 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.

Figure 1: Automatic data collection pipeline for processing parental provisioning videos of house sparrows on Lundy Island.

Parental investment videos were collected on Lundy Island UK, on days 7 and 11 post egg hatching. Videos were processed using open source software Deep Meerkat (Weinstein, 2018b), which applied convolutional neural networks to detect movement from wildlife monitoring videos, outputting frames with movement detected. Events were then defined by grouping movement frames that were less than 40 frames apart and at least 2 frames long. Lastly, events for each provisioning video were tallied to obtain an automatic feed rate as a proxy of parental provisioning.

Validation

After processing all videos using the automatic pipeline, we first tested for differences between the automatic and manual feed rates using Pearson's correlation test. To further compare different methods of video annotation, we randomly chose 8 videos to be used a test dataset. Using manual feed rates from previous annotation as the baseline value, we calculated three separate return rate measures by dividing the number of events detected using each method over the baseline number of events (*e.g.* if previous baseline found 10 feeding events and an alternative method found 7, the method has a return rate of 0.7). Firstly, we calculated the automatic pipeline return rate using the automatic feed rate obtained from the pipeline. Next, the machine learning (ML) assisted return rate

156 was calculated by manually reviewing the video clips extracted from the pipeline (Figure 1), to

157 manually filter out irrelevant non-feeding events. Lastly, we organized an undergraduate practical

158 session to annotate the same full-length videos to obtain an undergraduate return rate.

159

160 To test whether the extracted return rates were comparable to the baseline, we first did one sampled

161 t-tests to see whether each return rate was significantly different from 1 by setting the theoretical

162 mean (μ) to 1. We then compared the collected return rates with each other using pairwise t-tests

163 between each annotation method.

164

165 **Case study**

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

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

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

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

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

171 2019), we also ran the models separately using Day 7 and 11 rates to ensure results are consistent.

172

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

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

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

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

177 did not die in the nest and successfully fledged, while recruits were defined as fledglings that produced

178 at least one genetic offspring in their lifetime. To allow effect sizes to be compared, we z-transformed

179 the feed rates. To control for other effects that might affect the fitness metrics, we added the age of

180 the mother and father (see Wiebe, 2018), and hatch date (Days after April 1st) as fixed effects. Since

181 breeding success usually corresponds to peak food abundance (see Cresswell and Mccleery, 2003;

182   Lack, 1968), we also added a quadratic fixed-effect term for hatch date. Next, the population has

183   undergone routine cross-fostering, which is associated with increased survival (see Winney et al.,

184   2015), hence we added a fixed factor for fostered status (yes/no) in all models, however, note that

185   the feed rate we used were always from the social parent (parent that did the actual feeding). We

186   added the social parent IDs and the year as random effects to control for environmental effects (see

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

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

189   (Schroeder et al., 2012).

190

191   We ran all models specified above using the R packages MCMCglmm (Hadfield, 2010) on R version

192   3.6.1 (R Core Team, 2013). The posterior distributions and autocorrelations were checked following

193   Hadfield (2014) to ensure all fixed and random effects converged without violating any model

194   assumptions. We defined a parameter as statistically significant if the 95% credible interval did not

195   overlap with zero.

196

197　**Results**

198　**Qualitative Data**

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

200　(160 days) of recordings. We compared the extracted provisioning rates using a correlation (Figure 2),

201　revealing a significant positive correlation of 0.62 between automatic feed rates and manual feed

202　rates (Pearson's rank correlation; t_{781}=22.6, 95%CI: 0.58-0.66, p < 0.001).
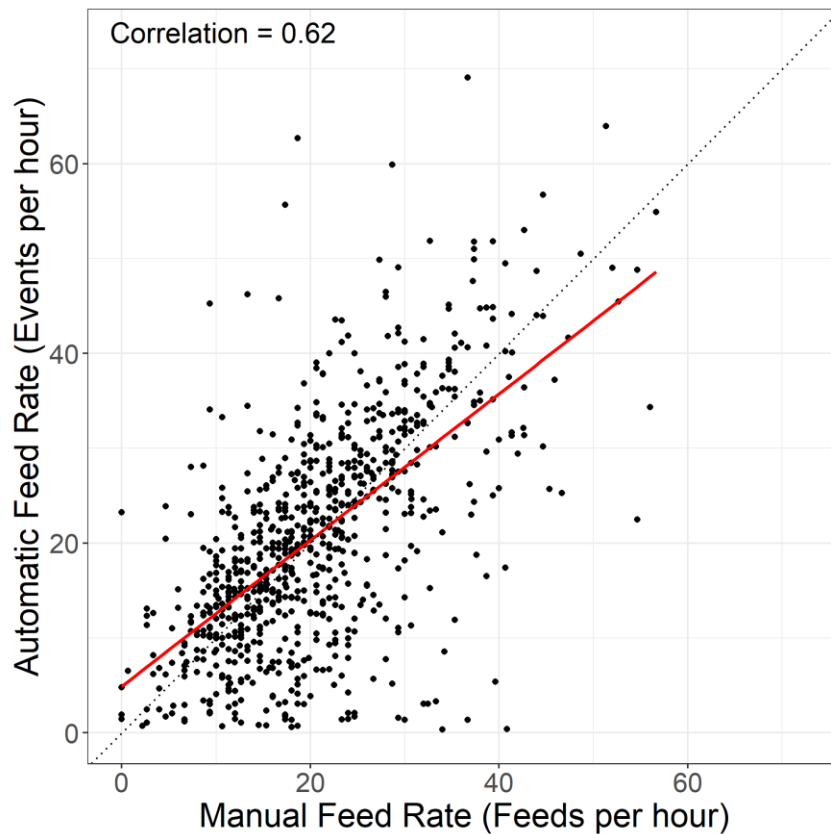
203

204　**Figure 2: Positive correlation between two measures of parental provisioning of house sparrows on**

205　**Lundy Island, UK.**

206　Manual feed rate refers to the number of feeding visits by both parents of a brood, measured manually

207　through video annotation. Automatic feed rate refers to the number of movement events detected

208　using the automatic video analysis pipeline divided by 2. The red line shows a significant positive

209　correlation between both measures and the dotted line shows a 1:1 line if both rates were equal.

210

211　**Validation**

212 Automatic pipeline return rate was significantly larger than 1 (Table 1) and significantly inflated

213 compared to other methods (Figure 3). However, once the extracted clips were manually reviewed to

214 eliminate non-feeding events to calculate the ML-assisted return rate, the rate is comparable with the

215 baseline (Table 1). The ML-assisted return rate is also not significantly different from undergraduate

216 annotation (Figure 3), both not significantly different from 1 (Table 1), comparable to the baseline

217 measure. On average, the ML-assisted approach took an average of 11.5 minutes per video (min 6.1

218 minutes, max 14.5 minutes).

219

220 **Table 1: Comparing event return rates of three data collection methods using a test dataset of 8**

221 **provisioning videos from house sparrows (*Passer domesticus*) on Lundy Island, UK.**

222 The return rate for each video was calculated by dividing the number of detected events obtained

223 using each method by the number obtained from previously annotated data. Test statistics were

224 obtained from a one-sampled t-test, with the theoretical mean (μ) set to 1. A) Videos were processed

225 by the pipeline to obtain an automatic feed rate. B) Videos were first processed by the pipeline, with

226 extracted clips further manually reviewed to filter out non-feeding events C) A cohort of

227 undergraduates annotated the same test videos, and the average number of annotated events for

228 each video were determined

229

Metrics	Mean	95% Confidence Interval	t	p-value
A) Automatic pipeline return rate	1.45	1.13-1.78	3.31	0.01
B) Machine learning assisted return rate	0.93	0.78 – 1.03	-1.86	0.10
C) Undergraduate return rate	1.03	0.96 – 1.11	1.13	0.30

230
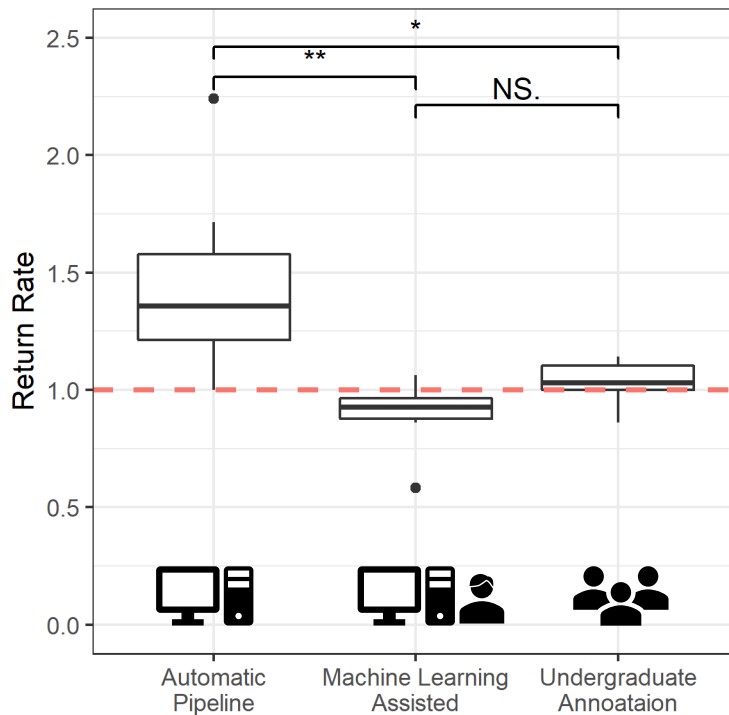
Figure 3: Comparison of return rates for three separate data collection methods on 8 parental provisioning videos of house sparrows (*Passer domesticus*) on Lundy Island, UK.

Return rates were calculated by dividing the number of feeding events each method detected over the number of events detected by manual annotation completed by past trained researchers. Automatic pipeline return rates were computed using the automatic feed rate collected from the video processing pipeline. Machine learning-assisted return rates were calculated by manually reviewing the short clips generated by the pipeline to remove non-feeding events. Undergraduate annotation return rates were collected by a cohort of undergraduate students annotating the original full-length videos. Red dotted line represents the reference level of 1, and brackets and labels represent the significance of t-test between all groups (N.S: not significant, *: $p<0.05$, **: $p<0.01$, ***: $p<0.001$).

Case study

The data that we used contained 658 unique broods, of 2116 individual chicks, of which 728 (34%) fledged and 328 (16%) were recruited. Automatic feed rates significantly predict an increase in the number of fledgelings and recruits for all models fitted (Figure 4; Supporting Information 1). We found

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

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

249
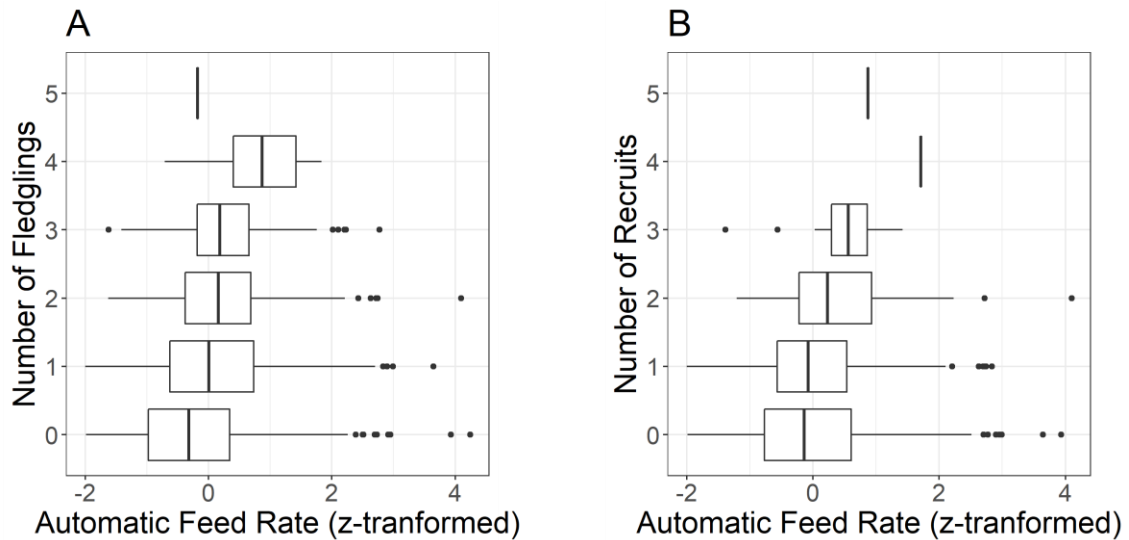
250

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

252 **increase in house sparrows (*Passer domesticus*) on Lundy Island, UK**

253 Automatic feed rates were obtained by processing provisioning videos using a machine learning

254 assisted pipeline and tallying the number of movement events detected then dividing by 2. Provision

255 rates were further z_tranformed to allow effect sizes to be comparable, with 0 representing the mean

256 and each unit representing one standard deviation away from the mean.

257

258

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. As such, we managed to process 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 8 test videos showed that the pipeline produced biased results, since the obtained automatic feed rate produced overinflated estimates compared to the baseline, likely due to over-detection of non-feeding events when birds are present within the video frame. However, the high 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 students was accurate, with mean event return rates higher than the reference baseline, possibly due

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

286   events due to misidentifications. We also showed that there were no significant difference between

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

288   approach. Given sex-specific hypotheses on parental investment are often tested in the population

289   (*e.g.* Schroeder et al., 2016), using the pipeline as a pre-processing step can be useful for collecting

290   additional sex or behavioural data within the feeding events. Further development using deep learning

291   and computer vision techniques can also classify the sex (see Ferreira et al., 2020) and behaviour (see

292   Conway et al., 2021; Ditria et al., 2021) automatically, which has potential to fully replace human

293   annotation in the future.

294

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

296   2021). Here, we presented an excellent case study of machine learning assisted data collection that

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

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

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

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

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

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

303   interesting hypotheses to be tested.

304

305   **Conclusion**

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

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

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

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

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

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

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

313 of laborious manual data collection.

314

315 **Significance Statement**

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

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

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

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

320

321 **Conflict of Interest Statement**

322 None

323 **Funding Statement**

328

329 **Acknowledgements**

332

333 **Author's Contributions**

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

335

336

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