



A high-throughput phenotyping pipeline for quinoa (*Chenopodium quinoa*) panicles using image analysis with convolutional neural networks

Journal:	<i>Plant Breeding</i>
Manuscript ID:	PLBR-24-OAr-260
Wiley - Manuscript type:	Original Article
Date Submitted by the Author:	29-Jun-2024
Complete List of Authors:	Lozano-Isla, Flavio; Universitat Hohenheim Institut fur Pflanzenzuchtung Saatgutforschung und Populationsgenetik; Universidad Nacional Toribio Rodriguez de Mendoza de Amazonas Kienbaum, Lydia; Universitat Hohenheim Institut fur Pflanzenzuchtung Saatgutforschung und Populationsgenetik Haussmann, Bettina I.G.; Universitat Hohenheim Institut fur Pflanzenzuchtung Saatgutforschung und Populationsgenetik; Deutsches Institut fur Tropische und Subtropische Landwirtschaft gGmbH Schmid, Karl; Universitat Hohenheim Institut fur Pflanzenzuchtung Saatgutforschung und Populationsgenetik
Keywords:	deep learning, genetic resources, genbank phenomics, ImageJ, plant breeding, image analysis
Note: The following files were submitted by the author for peer review, but cannot be converted to PDF. You must view these files (e.g. movies) online.	
Fig1.eps Fig3.eps Fig4.eps Fig5.eps ESM_1.html	

1
2
3 **1 A high-throughput phenotyping pipeline for quinoa (*Chenopodium quinoa*) panicles using image analysis with**
4 **2 convolutional neural networks**

5 3 Flavio Lozano-Isla^{1,2}, Lydia Kienbaum¹, Bettina I.G. Haussmann^{1,3}, Karl Schmid^{1*}

6 4 ¹ Institute of Plant Breeding, Seed Science and Population Genetics, University of Hohenheim, Stuttgart, Germany

7 5 ² Present affiliation: Instituto de Investigación para el Desarrollo Sustentable de Ceja de Selva (INDES-CES),
8 6 Universidad Nacional Toribio Rodríguez de Mendoza de Amazonas, Chachapoyas, Perú.

9 7 ³ Present affiliation: German Institute for Tropical and Subtropical Agriculture (DITSL GmbH), Witzenhausen,
10 8 Germany

11 9 *Corresponding author. Email: karl.schmid@uni-hohenheim.de

12 **10 ORCID IDs:**

13 11 Flavio Lozano-Isla: 0000-0002-0714-669X

14 12 Lydia Kienbaum: 0000-0003-0218-693X

15 13 Bettina I.G. Haussmann: 0000-0002-2360-6799

16 14 Karl Schmid: 0000-0001-5129-895X

17 15

1
2
3 **16 ABSTRACT**
4
5
6 **17** Quinoa is a grain crop with excellent nutritional properties which attracts global attention for its potential contribution
7 **18** to future food security in a changing climate. Grain yield is affected by panicle traits whose phenotypic analysis is
8 **19** time-consuming and error-prone because of their complex architecture, and automated image analysis is an efficient
9 **20** alternative. We designed a panicle phenotyping pipeline implemented in Python using Mask R-CNN for panicle
10 **21** segmentation and classification. We used it to analyze 5,151 images of quinoa panicles collected over three
11 **22** consecutive seasons from a breeding program in the Peruvian highlands. The pipeline follows a stage-wise approach,
12 **23** which first selects the optimal segmentation model and then another model that best classifies panicle shape. The best
13 **24** segmentation model achieved average precision (mAP) score of 83.16 and successfully extracted panicle length,
14 **25** width, area, and RGB values. The classification model achieved 95% prediction accuracy for panicle shape. Our
15 **26** pipeline enables accurate and cost-effective phenotyping of quinoa panicles. Using automated phenotyping based on
16 **27** deep learning, optimal panicle ideotypes can be selected in quinoa breeding.
17
18
19
20
21
22 **Keywords:** deep learning, genetic resources, genbank phenomics, ImageJ, plant breeding, image analysis
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

30 INTRODUCTION

31 Phenotyping is a key process in plant breeding programs aimed at developing new varieties of crops with higher yields.
32 However, phenotyping poses a bottleneck in the breeding process due to the labor-intensive nature of the task and
33 high cost. Manual phenotyping may result in reduced accuracy and prolonged timelines in breeding programs. Recent
34 advances in phenotyping technologies, such as the use of neural networks for image analysis, revolutionize plant
35 phenotyping by increasing precision, reducing labor intensity, and enabling the identification of new traits relevant to
36 crop breeding (Araus et al. 2018; Warman and Fowler 2021). This development of technology not only benefits major
37 crops but also minor and orphan crops because the limited resources available for the improvement of such crops can
38 be compensated by low-cost technological approaches for genetic and phenotypic analyses.

39 Quinoa (*Chenopodium quinoa* Willd.) is an example of an orphan crop. It originates from the Andean region, where
40 it has played a vital role as a staple food for small-scale farmers in the Andean highlands (Jacobsen et al. 2003; Hellin
41 and Higman 2005). For thousands of years, farmers have contributed to the domestication and selection of quinoa
42 varieties, resulting in a high genetic and phenotypic diversity (Bazile et al. 2016a; Patiranage et al. 2022). The grains
43 of quinoa are highly valued for their nutritional properties and serve as a rich source of macronutrients and energy
44 (Bhargava et al. 2006, Repo-Carrasco et al. 2003; Nowak et al. 2016; Chandra et al. 2018). Furthermore, quinoa is
45 remarkably resilient to abiotic stressors such as drought and salinity (reviewed by Grenfell-Shaw and Tester 2021).
46 Due to these benefits, quinoa cultivation has started to spread beyond its native area. Nevertheless, despite the
47 increasing number of countries growing quinoa, production volumes are small compared to the Andean region (Zurita-
48 Silva et al. 2014; Bazile et al. 2016). Although several breeding programs were established (Böndel and Schmid 2021),
49 modern quinoa varieties are still little improved compared to ancient landraces. The extensive phenotypic diversity of
50 quinoa in its center of origin reveals a substantial degree of genetic diversity, which likely results from local adaptation
51 and farmer selection (Bazile et al. 2016, Jarvis et al. 2017). This diversity manifests itself in various forms of
52 morphological and physiological variation, including panicle shape, seed color, leaf color (Bioversity International et
53 al. 2013), and varying degrees of tolerance to both biotic and abiotic stress factors. Recent efforts have aimed to
54 standardize quinoa evaluation and to characterize its diversity of traits for more effective application in breeding
55 programs (Stanschewski et al. 2021). However, the prevailing evaluation methods still largely rely on labor-intensive
56 and frequently imprecise manual phenotyping using visual scales and cards.

57 The inflorescence of quinoa is a panicle, on which individual flowers are arranged in groups on secondary axes
58 (Wrigley et al. 2015). The dimensions of the panicles, both in width and length influence grain yield, making them
59 target traits for selection in breeding programs (Benhabib et al. 2016a; Maliro et al. 2017; Santis et al. 2018). We
60 previously observed that the product of panicle length and panicle width was strongly correlated with yield and showed
61 a high heritability ($H^2=0.81$; Lozano-Isla et al. 2023). Two types of inflorescences are distinguished (Tapia et
62 al. 1979). The ‘glomerulate’ type panicles develop compact primary axillary internodes and elongated internodes of
63 flower clusters, which results in spherical inflorescences. The ‘amarantiform’ type develops elongated inflorescences
64 similar to amaranth species, and the finger-shaped partial inflorescences originate directly from the main axis.
65 Glomerulate inflorescences are considered to be wild type and are dominant over the amaranthiform type

(Gandarillas 1974), but the genetic basis of this variation is currently unknown. In the description of quinoa traits, a third type of panicle shape was defined as ‘intermediate’ (Bioversity International et al. 2013). Since grain yield is influenced by the size and type of the quinoa panicle (Craine et al. 2023), efficient phenotyping of these panicle traits is expected to improve the selection of superior genotypes and enhance genetic gain.

Computational image analysis is revolutionizing plant phenotyping because of its precision and throughput. In particular, neural networks emerge as a potent tool for phenotyping and characterizing crop diversity (Arya et al. 2022) because they frequently outperform traditional image analysis techniques (Liu and Wang 2020; Xie et al. 2020; Sabouri et al. 2021; Arya et al. 2022; Kang et al. 2023; Yu et al. 2023). The main objectives of image analysis are classification and image segmentation, which enable the measurement of traits such as size and shape. Additionally, it allows for the counting of features like fruits or color extraction (Ganesh et al. 2019; Zhou et al. 2019; Jia et al. 2020; Lee and Shin 2020). Among available algorithms for neural network-based image analysis, Mask R-Convolutional Neural Networks (Mask R-CNN; He et al. 2018), have been particularly successful, because they allow accurate pixel-wise mask prediction for objects. Mask R-CNN has found widespread applications in robotics (Jia et al. 2020), medicine (Anantharaman et al. 2018; Chiao et al. 2019), autonomous driving (Fujiyoshi et al. 2019), and plant science (Ganesh et al. 2019; Machefer et al. 2020; Jia et al. 2020; Kienbaum et al. 2021). Among classification models that identify and categorize objects within images, neural network architectures such as VGG16 (Simonyan and Zisserman 2015), InceptionV3 (Szegedy et al. 2016), and EfficientNetB0 (Tan and Le 2020) achieve state-of-the-art performance on multiple image classification tasks.

Considering the labor-intensive nature of phenotyping quinoa panicles under field conditions, we created a high-throughput pipeline for extracting phenotypic traits from images of quinoa panicles. This pipeline is based on our previous work (Kienbaum et al. 2021), which employed Mask R-CNN for the classification and segmentation of maize cob images. In the present study, we adapted and optimized this approach for image classification, specifically for differentiating between glomerulate and amaranthiform panicle types. We performed image segmentation to extract trait measurements to estimate quantitative-genetic parameters for yield components related to panicles. Our objective was to create a pipeline that achieves the same or higher level of accuracy as manual phenotyping but with significantly improved throughput, showcasing the effectiveness of state-of-the-art image analysis using Mask R-CNN in characterizing complex panicle traits. First, we describe the development of a deep learning model specifically designed for the classification and segmentation of quinoa panicles using Mask R-CNN. Subsequently, we demonstrate the application of this pipeline in estimating quantitative genetic parameters from a multi-location trial.

95

96 MATERIALS AND METHODS**97 Plant material**

98 The plant material used for the experiments was derived from six segregating populations of quinoa (*Chenopodium*
99 *quinoa* Willd.) originating from single crosses of landraces provided by the germplasm bank of the Universidad
100 Nacional del Altiplano, Puno, Peru (Lozano-Isla et al. 2023). The field trials were conducted in three successive
101 growing seasons from 2016 to 2019 as multi-location trials for line selection (“selection trials”), and variety
102 registration (“registration trials”) trials in the Peruvian Highlands (Table 1). Images from the 2021-2022 season for
103 production and seed increase (“production trials”) were used for the validation set using scales in the images.

104 Panicle images

105 Pictures were taken during flowering at stage 69 of the BBCH scale (Sosa-Zuniga et al. 2017). A representative panicle
106 for each experimental unit (breeding lines grown in experimental plots) was selected for photographing under field
107 conditions in front of a blue background with different light conditions (Fig. 1, Figure S1). Images were taken with
108 different cameras and resolutions (Table 1), resulting in an image collection exhibiting a high heterogeneity (Fig. 1a).
109 We excluded images with panicles that showed bird damage, were blurred, or were overly dry from the analysis
110 (Fig. 1b).

111 Development of a Mask R-CNN mode for image segmentation

112 We found it challenging to obtain a single Mask R-CNN model suitable for both panicle segmentation and
113 classification. Mask R-CNN is simple to train and extends Faster R-CNN by adding a branch for predicting an object
114 mask in parallel with the existing branch for bounding box recognition (Bharati and Pramanik 2020). Consequently,
115 we adopted a stagewise approach, selecting separate models for each task. A pipeline was implemented, connecting
116 the best models of each stage. First, we developed models for image segmentation by testing 16 models and selecting
117 the best one (He et al. 2018) using the following procedure.

118 Sampling and annotation

119 To ensure the selection of an unbiased training dataset, we created two random samples. The first sample involved
120 selecting 25% from each experiment conducted in different years and locations (Table 1) making a total of ca. 1,300
121 images. A random sample of 320 images was taken from this initial pool of images. The sample was divided into
122 training and validation sets with 238 and 82 images respectively. The number of training images was determined based
123 on a previous study of maize cob segmentation (Kienbaum et al. 2021). The images were annotated using VGG Image
124 Annotator version 2.0.8 (Dutta and Zisserman 2019). For selection and registration field trials, only the panicle class
125 was annotated (Table 1). For the production field trial, seven classes were determined for the images: panicle, scale,
126 label, barcode, and each channel from RGB spectra (Figure S2).

1
2
3 127 ***Model training***
4

5 128 We utilized the Mask R-CNN framework (He et al. 2018) and fitted it with the Resnet 101 backend to the seven
6 classes. Model training was conducted on four parallel Tesla K80 GPUs on the BinAC high-performance computing
7 cluster at the University of Tübingen. The Mask R-CNN training parameters were partly selected based on previous
8 work in image analysis for maize cobs (Kienbaum et al. 2021), with further variations and improvements implemented
9 as needed. Quinoa panicles exhibit greater detail and diversity in their outline and shape compared to maize cobs.
10
11 132 Therefore, we varied parameters such as mask resolution, loss weight, and training configuration (heads.m) to
12 investigate potential differences in the final model performance. This led to training 16 distinct Mask R-CNN models
13 with a learning rate of 1e-4 over 200 epochs, each with different parameters (Table 2).
14
15

16 136 In Mask R-CNN, the mask resolution determines the size of the predicted masks for each object instance. A higher
17 resolution yields a more detailed delineation of object boundaries but increases computational time. We adjusted this
18 parameter from the standard resolution of 28x28 pixels to an enhanced resolution of 56x56. This change required a
19 minor modification in the Mask R-CNN architecture, which involved adding an extra convolutional layer at the end
20 of the mask branch. The aim was to achieve a more precise panicle mask.
21
22

23 140
24 141 Model optimization was based on the loss weight parameter. It ranged from the standard loss weight of 1 for each
25 mask and class (mask01-class01) to an emphasis on mask optimization with a mask loss weight of 10 and a class loss
26 weight of 1 (mask10-class01). We also experimented with classification optimization, adjusting the class loss weight
27 to 10 and mask loss weight to 1 (mask01-class10). The parameter heads.m indicated the training configuration, which
28 either involved training all layers of the Resnet-101 architecture or fine-tuning only the head layers while freezing all
29 other layers **(Table 2).
30
31

32 146
33 147 The performance of the segmentation model in detecting panicles was assessed using the Intersection over Union
34 (IoU) score, also known as the Jaccard index (Jaccard 1901). This metric is widely used in evaluating the performance
35 of object detection. The effectiveness of our trained models was measured using the mean average precision (mAP)
36 over different IoU thresholds ranging from 50 to 95% in 5% increments (AP@[IoU = 0.50:0.95]; Kienbaum et
37 al. 2021).
38
39

40 151
41 152 ***Model development for image classification***
42

43 153 In the second stage, three deep learning architectures were implemented and tested for model classification using
44 VGG16, InceptionV3, and EfficientNetB0 to differentiate between the two panicle types. VGG16 has the advantage
45 of accurate identification and performs better on large-scale data sets and complex background recognition tasks
46 (Wang 2020). InceptionV3 is characterized by its multi-level feature extraction and factorization techniques, achieving
47 a balance between accuracy and computational efficiency (Li et al. 2021). The efficientNetB0 compound scaling
48 approach offers high performance while maintaining reduced computational complexity, which is useful for real-time
49 applications and resource-constrained environments (Ramamurthy et al. 2023).
50
51

160 *Sampling*

161 A total of 320 panicle masks from the segmentation pipeline and pictures from the experiments were randomly selected
162 and divided into two groups: training and validation. Each group contained two panicle classes: glomerulate and
163 amaranthiform. The training group included 110 images of the amaranthiform and 142 of the glomerulate shape, whereas
164 the validation group included 30 and 36 images of the same respective panicle classes. The imbalance in the number
165 of panicles in the datasets primarily reflects the predominance of glomerulate panicles over amaranthiform panicles.

166 *Model training*

167 To classify images into two classes of panicle shapes, namely amaranthiform and glomerulate, we implemented 12
168 models using various combinations of convolutional neural network architectures, dense layers, and activation
169 functions (Table 3). Each model was constructed using a specific neural network architecture such as VGG16,
170 InceptionV3, or EfficientNetB0, and included two dense layers. The first layer employed a ReLU (Rectified Linear
171 Unit) activation function and utilized either 128 or 1024 dense layers. The choice to vary the number of neurons was
172 influenced by computational resource limitations and the need for efficient model utilization. The second layer used
173 either a sigmoid or softmax output activation function for image classification (Maharjan et al. 2020). We employed
174 a standard image augmentation technique, commonly used to artificially increase the size and diversity of the image
175 dataset, for a more robust model. This involved horizontal flips, random crops, random Gaussian blur, varied contrast,
176 brightness, zoom, translation, and rotation using the *imgaug* library in Python (Jung 2022). The models were executed
177 under three replications for 200 epochs (Table 3). The ModelCheckpoint function was used to automatically save the
178 model exhibiting the highest performance, based on the lowest validation loss. The prediction accuracy (%) for the
179 two panicle classes, amaranthiform and glomerulate, was evaluated considering the three factors: neural network
180 architectures, dense layers, and activation functions.

181 *Pipeline for quinoa panicle image analysis*

182 The best segmentation model, selected based on the mAP score, was then used to export the pixel-wise mask of
183 panicles from each image and to extract the phenotypic traits: panicle, scale, label, barcode, and RGB values by
184 channel (Fig. 2c-d). The panicle masks were submitted to the best classification model to identify the panicle shape
185 as either amaranthiform or glomerulate (Fig. 2e-f).

186 For the first image dataset, selection, and registration trials, the length, width, area in pixels, and RGB values were
187 extracted. The barcode and image scale were included for the second image dataset, the production trial. Since the
188 images of the first dataset did not contain scales that would allow us to calculate absolute measurements. We calculated
189 two indices related to grain yield (Lozano-Isla et al. 2023) based on panicle length and width (i.e. panicle width/length
190 and length/width) with the aim of determining if there are differences in genetic parameters such as heritability. The
191 pipeline ran on a workstation with 64GB RAM and 8 CPU kernels with a total of 16 threads.

192 Deep learning pipeline versus manual annotation

193 Fifteen panicle images from the selection and registration field trials (Table 1) were randomly selected and manually
194 measured using ImageJ (Schneider et al. 2012) for comparison with the results from the segmentation model. The
195 measures extracted using ImageJ were panicle area, width, and length (all measured in pixels), and mean and standard
196 deviation of each RGB channel (Figure S3). The results were compared using Pearson correlation analysis. To extract
197 the RGB value from the mask, the RGB Measure plugin was installed from <https://imagej.nih.gov/ij/plugins/rbg->
198 measure.html.

199 Quantitative genetic analysis of panicles in multilocation trials

Due to the imbalance in the number of entries between years (Table 1), the indices were analyzed using a stagewise approach (Schmidt et al. 2019; Buntaran et al. 2020) under multi-location trials with three locations. In stage 1, a linear mixed model with lines as the fixed effect was used for each experiment under the lattice design (Zystro et al. 2018). This was done to estimate the Best Linear Unbiased Estimators (BLUEs) using the following model:

$$\gamma_{ijk} = \mu + g_i + rep_j + block_{jk} + plot_{ijk}$$

where γ_{ijk} is the response variable of the i^{th} genotype in the k^{th} block of the j^{th} replicate, μ is the first-stage intercept, g_i is the effect for the i^{th} genotype in the first stage, rep_j is the effect of the j^{th} replicate, $block_{jk}$ is the effect of the k^{th} incomplete block of the j^{th} replicate, and $plot_{ijk}$ is the plot error effect corresponding to γ_{ijk} .

208 In stage 2, a linear mixed model for line by environment interaction, where the lines were treated as fixed effects. This
209 was done to calculate the Best Linear Unbiased Predictors (BLUPs) using the following model:

$$\bar{\gamma}_{im} = \mu + g_i + l_m + gl_{im} + \bar{e}_{im}$$

where $\bar{\gamma}_{ihm}$ is the adjusted mean of the i^{th} genotype in the m^{th} location obtained in the first stage, μ is the intercept, l_m is the main effect for the m^{th} location, g_i is the main effect of the i^{th} genotype, gl_{im} is the im^{th} genotype x location interaction effect and \bar{e}_{im} is the error of the mean γ_{im} obtained in the first stage.

214 Statistical analysis

215 To conduct statistical analyses and produce graphs, we used the statistical programming language R version 4.4.0 (R
216 Core Team 2020), while image analysis was performed using Python version 3.7.

The image analysis models were compared with different parameters (i.e., segmentation and classification), and each model with its respective replicas was subjected to a variance analysis with Anova(type = “III”) in the *car* package (Fox et al. 2023). The factors that showed significance were subjected to a means comparison analysis using the Tukey comparison test ($p < 0.05$) implemented in the package *emmeans* (Lenth et al. 2023) and *multcomp* (Hothorn et al. 2023).

1
2
3 222 For the multi-environment analyses, broad-sense heritability (i.e., Cullis approach), variance components, Best Linear
4 223 Unbiased Estimators (BLUEs), and Best Linear Unbiased Predictors (BLUPs) were estimated based on the ‘H2cal()’
5 224 function of the ‘inti’ package (Lozano-Isla 2021). This function uses a linear mixed model for both random and fixed
6 225 effects for genotypes, based on the *lme4* package (Bates et al. 2020). Outlier removal for multi-location trials was
7 226 based on method 4, Bonferroni-Holm, using re-scaled median absolute deviation for standardizing residuals as
8 227 described in Bernal-Vasquez et al. (2016) and implemented in the same function.
9
10
11

12 228 A Pearson correlation plot was used to compare the trait predictions made with Mask R-CNN and ImageJ and was
13 229 performed using the *psych* R package (Revelle 2021). The plots were produced with the package *ggplot2* (Wickham
14 230 et al. 2023) and *ggsicle* (Landis 2022). Code and reproducible data analysis were implemented under Quarto, an open-
15 231 source scientific and technical publishing system (Allaire et al. 2022, Supplementary File 1).
16
17

18 232
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

233 RESULTS**234 Training models for segmentation and classification**

235 To implement a pipeline for quinoa panicle analysis, a comparative analysis of various segmentation and classification
236 models was conducted. The selection of the best segmentation model involved the assessment of 16 Mask R-CNN
237 models, each tested with a combination of parameters, mask resolution, loss weight, and neural network. The accuracy
238 of the segmentation models was evaluated using the mean average precision (mAP) metric (Table 2). The selection of
239 the best classification model entailed the comparison of 12 classification models that were developed through different
240 combinations of neural network architectures, dense layers, and activation functions. The accuracy of the classification
241 models was assessed based on the predictive performance for two classes (i.e. amaranthiform and glomerulate, Table 3).

242 The segmentation models showed statistical significance in the interaction of the loss weight, mask resolution, and
243 neural network ($p = 0.01$, Table 2, Supplementary Table 1). The mAP scores ranged from 0.68 to 0.8, with a coefficient
244 of variation of 3.65%. The model segmentation-09 achieved the highest mAP score of 0.8, while the model
245 segmentation-08 obtained the lowest score at 0.69 (Table 2). The segmentation model selected for implementation in
246 the pipeline was chosen based on the replica that achieved the highest mAP with 0.83 (Supplementary Table 2).

247 For classification, the prediction accuracy results for the glomerulate and amaranthiform classes ranged from 0.85 to
248 0.92. No significant differences were observed among the factors studied in the classification models (Table 3,
249 Supplementary Table 3). Therefore, we selected model classification-08 with the highest accuracy of 0.95 in the
250 repeated model evaluations for further analyses (Supplementary Table 4).

251 Pipeline for quinoa panicle phenotyping

252 We implemented a Python pipeline to extract phenotypic traits from quinoa panicles, leveraging the best model for
253 segmentation and classification analysis (Fig. 2). To segment panicle images, we used the model segmentation-09
254 (Fig. 2a-d). As certain images featured multiple panicles, we generated individual masks for each panicle.
255 Subsequently, these masks were employed in the classification process. The segmentation model's pixel-wise masking
256 of each panicle allowed us to extract panicle length, width, area, and mean RGB color values for each channel. The
257 segmentation of panicles was followed by the input of these segmented structures into the classification model.
258 Subsequently, the model categorized the images based on one of two panicle shapes: glomerulate or amaranthiform
259 (Fig. 2e-f). We combined the results from the segmentation and classification pipeline (Fig. 2g) and used them to
260 calculate quantitative-genetic parameters such as variance components, BLUEs, BLUPs, and heritability. The results
261 generated by this pipeline can be utilized to carry on association studies, such as Genome-Wide Association Studies
262 (GWAS), or Quantitative Trait Locus (QTL) mapping (Fig. 2h).

263

1
2
3 **264 ImageJ versus deep learning pipeline**

4
5 **265** To assess the performance of the deep learning pipeline, we conducted a regression analysis comparing manual trait
6 extraction using ImageJ software with the results obtained from our model pipeline (Fig. 3).

7 **267** We selected a total of 60 images, with fifteen images from each experiment (Table 1), and manually annotated them
8 with ImageJ (Figure S1) for panicle length, width, and area. The regression (R^2) for panicle length was 0.93 ($r = 0.96$),
9 for width 0.92 ($r = 0.96$), and area 0.99 ($r = 0.99$) showing a high correlation (Fig. 3a-c). Panicle distribution for width
10 and area shows similar distributions between manual annotation and model pipeline.

11 **271** The classification model achieved an accuracy of 97.6%. The sensitivity was 96.7%, representing the proportion of
12 panicles correctly identified as panicle-positive by the model (Fig. 3e). The specificity of the model was 100%,
13 indicating the proportion of individuals without panicles correctly identified as panicle-negative (Fig. 3f). The model's
14 precision, or the proportion of positive predicted values, was 100%.

15
16 **275 Scaled images and deep learning pipeline**

17
18 **276** To evaluate the efficiency of the pipeline under field conditions, we updated the segmentation model to include a
19 dataset with 106 images with scale and QR codes (Fig. 4). We took photos during the seed production campaign using
20 a 10 cm scale (Table 1, Fig. 4f).

21
22 **279** The regression analysis between the manual measurement and the model prediction did not show a correlation for the
23 panicle length ($R^2 = <0.01$, Fig. 4a). In contrast, panicle width showed a significant correlation ($r = 0.73$, $R^2 = 0.54$,
24 Fig. 4b). A medium correlation was presented for the panicle indices. For width/length ($r = 0.6$, $R^2 = 0.36$ Fig. 4c) and
25 width/length ($r = 0.56$, $R^2 = 0.31$, Fig. 4d). The classification model achieved an accuracy of 98.1%. The sensitivity
26 was 100% and the model's specificity was 94.1% (Fig. 4e).

27
28 **284 Quantitative-genetic analysis from the multi-location trials**

29
30 **285** To assess the effectiveness of the pipeline within the framework of breeding programs. We computed the quantitative-
31 genetic parameters by utilizing the extracted phenotypic values of quinoa panicles through the pipeline (Fig. 2). We
32 conducted a stagewise analysis under a multi-location trial for the registration trials (Table 1). In the first stage, we
33 calculated the adjusted means. In the second stage, we performed a model selection to calculate the BLUPs, choosing
34 the model with the lowest AIC (Supplementary Table 5). In the absence of scales in the images from the registration
35 trials, we proceeded to estimate the broad-sense heritability for two indices, specifically derived from the length and
36 width measurements of the panicles (Fig. 5).

37
38 **292** The predominant panicle shape was glomerulate (Fig. 5a). Its frequency was 82% among F7 genotypes (Illpa site),
39 while in the F8 generation, its frequency was 84% (Camacani) and 63 % (Illpa). ANOVA shows a significant
40 interaction between panicle shape and its length-to-width ratio ($p < 0.001$). Glomerulate panicles were found to be
41 longer than they were in width, whereas in the case of amaranthiform panicles, a tendency to be wider and less long
42 (Fig. 5b).

1
2
3 297 For the quantitative genetic parameters, the highest heritability was observed for the ratio of panicle width to length
4 298 ($H^2 = 0.61$), while panicle length to width showed a heritability of 0.53 (Fig. 5c). The evaluated traits revealed
5 299 significant genetic variance, with values of 48.8 and 58.3 for the ratios of length to width and width to length,
6 300 respectively. The panicle length-to-width ratio exhibited a better normal distribution compared to the width-to-length
7 301 ratio (Fig. 5d).

8
9 302 Panicle shape is a bimodal trait (i.e., amaranthiform and glomerulate), and it does not exhibit a normal distribution
10 303 (Fig. 5d). To compare the heritability using linear and generalized mixed models, we analyzed the results with a linear
11 304 mixed model (lmer) and a generalized linear mixed model with a binomial distribution (glmer; Supplementary Table
12 305 8). For lmer, the standard heritability was 0.66, while for the glmer model it was 0.69. However, as we cannot calculate
13 306 the error variance for the glmer model, this results in a 4% difference between the outcomes (Supplementary Table
14 307 6). Based on this result, we decided to proceed with the analysis using lmer.

15
16 308 Using the mixed linear model, we observed high repeatability for panicle shape in the F7 season in Illpa when
17 309 experimental units had replicates (Supplementary Table 7). In the second phase of the analysis, using the dataset from
18 310 the experiments in Illpa and Camacani for the F8 season, we found that heritability decreased to 0.24 when
19 311 experimental units did not have replicates (Fig. 5c, Supplementary Table 7). However, the standard heritability for
20 312 panicle type is 0.63 ($H^2.s$, Supplementary Table 8).

21
22 313

314 DISCUSSION

315 Panicle traits are important components for increasing yield in quinoa. However, field-based panicle phenotyping is
316 time-consuming and labor-intensive, limiting its use in breeding programs and large-scale genetic studies such as QTL
317 mapping or GWAS. To address this limitation, we aimed to develop a pipeline for high-throughput phenotyping of
318 quinoa panicles using deep learning-based image analysis. We built a pipeline that selects the best models using a
319 stagewise approach. In the first stage, we tested 16 different models to identify the most suitable one for segmenting
320 panicle images. In the second stage, we implemented 12 classification models to categorize images into two panicle
321 shapes: amaranthiform and glomerulate. To verify the accuracy of the pipeline, we compared the output of tested
322 models with manual evaluations using ImageJ. The pipeline output demonstrated a high correlation with manual
323 evaluations at a pixel-based scale for segmentation and a high prediction accuracy for quinoa panicle shape. We
324 applied the complete pipeline to calculate quantitative-genetic parameters in multi-location trials, exemplifying its
325 implementation for breeding programs.

326 Pipeline to panicle image analysis in quinoa

327 Despite training various Mask R-CNN models with different parameters, our pipeline was not able to identify a single
328 model that performed best in both classification and segmentation tasks simultaneously. Therefore, we decided to
329 implement a stagewise approach selecting a separate model for each task for panicle phenotyping.

330 In a previous study on maize cob image analysis (Kienbaum et al. 2021), a single Mask R-CNN model reliably
331 predicted the classes and masks of maize cobs and ruler elements. The features of maize cobs and rulers were simpler
332 and more distinguishable than the different quinoa panicles with the presence of leaves, similar colors between
333 structures (i.e., stem, leaves, and panicle), and in some cases, the panicles not exhibiting compact density. The need
334 to choose two models probably relies on the difficulty of optimizing both for the challenging classification between
335 the panicle shapes and for accurately segmenting the detailed panicle traits with the masks. We selected a segmentation
336 model that had a score of 83.16 mAP with two classes, a similar result to what Kienbaum et al., (2021) found with
337 two classes at 87.7 mAP in maize (Kienbaum et al. 2021), and with two classes at 89.85 mAP in strawberries (Yu et
338 al. 2019).

339 Effective detection using a simple image processing method can be challenging under field conditions due to various
340 environmental such as background color, image position, shadow, and image rotation. Lee and Shin (2020) managed
341 to detect potato shapes under field conditions and calculated parameters based on pixels using Mask R-CNN. We
342 adopted a similar approach in this study to detect and analyze quinoa panicles under field conditions, showing that
343 Mask R-CNN is a suitable method to reduce the tedious and costly labor associated with traditional approaches. Zhou
344 et al. (2019) analyzed 1,064 panicles from 272 genotypes of sorghum and extracted the area, volume, and panicle
345 length and width, showing a high correlation with manual annotation. We obtained similar results with our approach
346 when we compared it with manual annotation (Fig. 3).

1
2
3 347 Training a model for image segmentation, compared to image classification, requires annotation of both class and
4 object location information in images for model training, which can be a time-consuming task. In our dataset, many
5 photos did not contain complete panicles, because they were taken under field conditions by students (and not
6 experienced photographers), suggesting the need for adjustments in image acquisition protocols and better training of
7 personnel taking photographs. Additionally, the images from selection and registration trials did not have any scale,
8 so we had to rely on panicle indices based on relative rather than trait values for our analysis. For panicle classification,
9 we did not include intermediate panicle shapes, which are challenging to differentiate, even with visual scoring (Craine
10 et al. 2023).
11
12
13
14

15 355 **Quantitative-genetic analysis of multi-location trials**
16
17

18 356 The primary objective of our pipeline was to facilitate high-quality phenotyping of quinoa panicles, enabling reliable
19 357 estimation of quantitative-genetic parameters such as variance components and heritability. This information
20 358 facilitates the prediction of responses to selection in breeding programs (Visscher et al. 2008).

21 359 To minimize measurement errors and improve estimates, it is necessary to investigate a large number of genotypes.
22 360 For these reasons, automatic and semi-automatic methods have been proposed for analyzing panicles in rice and
23 361 sorghum (Zhou et al. 2019; Kong and Chen 2021). In both cases, the panicles were removed from the plants and
24 362 analyzed under controlled conditions, leading to increased labor and cost. Our pipeline robustly analyzes images taken
25 363 during crop development in a non-destructive way under variable field conditions, differing for example for light,
26 364 angle, shadows, background color, and photographic devices, all of which create variation in image quality.

27 365 Phenotyping and quantifying panicle traits are crucial as they correlate with the yield of quinoa. Several studies have
28 366 found a positive correlation between grain yield and both panicle length and panicle width (Benlhabib et al. 2016b;
29 367 Maliro et al. 2017; Santis et al. 2018). We used the ratio between panicle Width and Length, which exhibits a 0.86
30 368 heritability at maturity (Lozano et al. 2022). Efficient estimation of quantitative-genetic parameters for panicle traits
31 369 could enable more efficient selection and increase the selection gain for yield in future breeding populations.

32 370 According to quinoa descriptors, there are three different types of quinoa panicles: amaranthiform, intermediate, and
33 371 glomerulate (Bioversity International, 2013). Intermediate panicle types are challenging to distinguish, even for the
34 372 human eye. For this reason, we classified the panicles into the two most distinct types, glomerulate and amaranthiform.
35 373 Future analysis also might include panicle density, described as lax versus compact, as it could be a valuable trait for
36 374 breeding programs and is subject to environmental effects (Manjarres-Hernández et al. 2021). The selection of
37 375 amaranthiform panicles can promote increased airflow throughout the quinoa inflorescences leading to a reduced
38 376 incidence of pests and diseases. The integration of panicle shape and density has the potential to enhance the panicle
39 377 ideotype in quinoa.

40 378
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

1
2
3 380 A significant limitation of our analysis is the absence of a scale in the images taken in the selection and registration
4 field trials. This absence prevented us from using absolute measurements, such as panicle length and width in
5 centimeters, for quantitative genetic analysis. Instead, we had to depend on a relative index from the ratio between
6 panicle length and width. We included a few images from the seed production trial and updated the pipeline with
7 images that include barcode and scale. The information on the pictures was extracted from the barcodes and the scales
8 to extract the panicle traits in centimeters. This implementation significantly streamlines image processing and
9 subsequent analysis. The updated model using scales was not able to predict the real length with accuracy, maybe due
10 to two main reasons, (i) the few numbers of images included in the model training, and (ii) the scale was not distributed
11 around the picture frame during the picture acquisition. Despite this, our pipeline is now equipped for large-scale
12 analysis of panicle images that include panicles, barcodes, and scales.
13
14

15 390 However, taking the panicle images correctly in the field could be time-consuming than scoring the two shapes
16 visually than taking images and running the pipeline. Nevertheless, parameters such as area, color, and leaflet in the
17 panicles could be difficult to score visually. Color is typically considered a qualitative trait, the ability to extract RGB
18 values could prove useful for genetic analysis, as demonstrated for seed color in peanuts by Zhang et al. (2020). The
19 use of high-throughput phenotyping using deep learning opens new opportunities for the characterization and
20 evaluation of the diversity in large numbers of panicle images in quinoa. Implementation of this approach in plant
21 breeding programs can increase phenotyping efficiency and improve the selection of future quinoa varieties with better
22 panicle traits.
23
24

25 398 In conclusion, we developed a pipeline to be directly applicable to any set of panicle images without requiring prior
26 model training. It is user-friendly, easy to execute, and does not demand expensive computational resources such as
27 GPUs. For this reason, it is highly suitable for quinoa breeders with limited resources.
28
29

30 401
31
32 402 **STATEMENTS AND DECLARATIONS**
33
34 403 **Acknowledgments**
35
36 404 We acknowledge support from the High Performance and Cloud Computing Group at the Zentrum für
37 Datenverarbeitung of the University of Tübingen, the state of Baden-Württemberg through bwHPC, and the German
38 Research Foundation (DFG) through grant no INST 37/935-1 FUGG. We thank Emilia Koch for the image annotation.
39 Thanks to Paul, Jakob, Jose David, Miguel, Edwin, and Blander who helped take the images in the field trials. Many
40 thanks to Felix Bartusch from the University of Tübingen for his support in utilizing the BinAC high-performance
41 computing infrastructure.
42
43

44 410 **Data availability statement**
45
46 411 The code used for analysis and images for annotation is available at the GitHub repository:
47
48 412 https://github.com/Flavjack/quinoa_panicle_phenotyping
49
50

51 413 **Funding statement**
52
53
54
55
56
57
58
59
60

1
2
3 414 This work was funded by a capacity development program of KWS SAAT SE & CO KGaA grant to the Universidad
4 415 Nacional del Altiplano Puno and the University of Hohenheim.
5
6 416 **Conflict of interest statement**
7
8 417 The authors declare that the research was conducted without any commercial or financial relationships that could be
9 418 a potential conflict of interest.
10
11 419 **Author contributions**
12
13 420 Conception and design of the study by FLI, LK, BIGH, KS. FLI and LK performed material preparation, data
14 421 collection, and analysis. FLI and LK wrote the first draft of the manuscript. All authors commented on previous
15 422 versions of the manuscript. All authors read and approved the final manuscript.
16
17
18 423
19
20
21 424 **REFERENCES**
22
23 425 Allaire JJ, Teague C, Scheidegger C, et al (2022) Quarto: open-source scientific and technical publishing system built
24 426 on Pandoc
25
26 427 Anantharaman R, Velazquez M, Lee Y (2018) Utilizing Mask R-CNN for Detection and Segmentation of Oral
27 428 Diseases. In: 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). pp 2197–2204
28
29 429 Araus JL, Kefauver SC, Zaman-Allah M, et al (2018) Translating High-Throughput Phenotyping into Genetic Gain.
30 430 Trends Plant Sci 23:451–466. <https://doi.org/10.1016/j.tplants.2018.02.001>
31
32
33 431 Arya S, Sandhu KS, Singh J, kumar S (2022) Deep learning: as the new frontier in high-throughput plant phenotyping.
34 432 Euphytica 218:47. <https://doi.org/10.1007/s10681-022-02992-3>
35
36 433 Bates D, Maechler M, Bolker B, Walker S (2020) lme4: Linear mixed-effects models using “eigen” and S4
37
38 434 Bazile D, Jacobsen S-E, Verniau A (2016a) The Global Expansion of Quinoa: Trends and Limits. Front Plant Sci 7:
39 435 <https://doi.org/10.3389/fpls.2016.00622>
40
41 436 Bazile D, Pulvento C, Verniau A, et al (2016b) Worldwide Evaluations of Quinoa: Preliminary Results from Post
42 437 International Year of Quinoa FAO Projects in Nine Countries. Front Plant Sci 7:
43 438 <https://doi.org/10.3389/fpls.2016.00850>
44
45 439 Benhabib O, Boujartani N, Maughan PJ, et al (2016a) Elevated Genetic Diversity in an F2:6 Population of Quinoa
46 440 (*Chenopodium quinoa*) Developed through an Inter-ecotype Cross. Front Plant Sci 7:
47 441 <https://doi.org/10.3389/fpls.2016.01222>
48
49 442 Benhabib O, Boujartani N, Maughan PJ, et al (2016b) Elevated Genetic Diversity in an F2:6 Population of Quinoa
50 443 (*Chenopodium quinoa*) Developed through an Inter-ecotype Cross. Front Plant Sci 7:
51
52
53
54
55
56
57
58
59
60

- 1
2
3 444 Bernal-Vasquez A-M, Utz H-F, Piepho H-P (2016) Outlier detection methods for generalized lattices: a case study on
4 445 the transition from ANOVA to REML. *Theor Appl Genet* 129:787–804. <https://doi.org/10.1007/s00122-016-2666-6>
- 5 446 Bharati P, Pramanik A (2020) Deep Learning Techniques—R-CNN to Mask R-CNN: A Survey. In: Das AK, Nayak
6 447 J, Naik B, et al. (eds) Computational Intelligence in Pattern Recognition. Springer, Singapore, pp 657–668
- 7
8
9
10 448 Bioversity International, Fundación para la Promoción e Investigación de Productos Andinos, Instituto Nacional de
11 449 Innovación Agropecuaria y Forestal, et al (2013) Descriptors for quinoa (*Chenopodium quinoa* Willd) and wild
12 450 relatives. Bioversity International
- 13
14 451 Böndel KB, Schmid KJ (2021) Quinoa Diversity and Its Implications for Breeding. In: Schmöckel SM (ed) The
15 452 Quinoa Genome. Springer International Publishing, Cham, pp 107–118
- 16
17 453 Buntaran H, Piepho H-P, Schmidt P, et al (2020) Cross-validation of stagewise mixed-model analysis of Swedish
18 454 variety trials with winter wheat and spring barley. *Crop Sci* 60:2221–2240. <https://doi.org/10.1002/csc2.20177>
- 19
20 455 Chiao J-Y, Chen K-Y, Liao KY-K, et al (2019) Detection and classification the breast tumors using mask R-CNN on
21 456 sonograms. *Medicine (Baltimore)* 98:e15200. <https://doi.org/10.1097/MD.00000000000015200>
- 22
23 457 Craine EB, Davies A, Packer D, et al (2023) A comprehensive characterization of agronomic and end-use quality
24 458 phenotypes across a quinoa world core collection. *Front Plant Sci* 14:1101547.
25 459 <https://doi.org/10.3389/fpls.2023.1101547>
- 26
27 460 Dutta A, Zisserman A (2019) The VIA Annotation Software for Images, Audio and Video. In: Proceedings of the 27th
28 461 ACM International Conference on Multimedia. Association for Computing Machinery, New York, NY, USA, pp
29 462 2276–2279
- 30
31 463 Fox J, Weisberg S, Price B, et al (2023) car: Companion to Applied Regression
- 32
33 464 Fujiyoshi H, Hirakawa T, Yamashita T (2019) Deep learning-based image recognition for autonomous driving. IATSS
34 465 Res 43:244–252. <https://doi.org/10.1016/j.iatssr.2019.11.008>
- 35
36 466 Gandarillas H (1974) Genética y origen de la quinua. Instituto Nacional del Trigo, La Paz, Bolivia
- 37
38 467 Ganesh P, Volle K, Burks TF, Mehta SS (2019) Deep Orange: Mask R-CNN based Orange Detection and
39 468 Segmentation. IFAC-Pap 52:70–75. <https://doi.org/10.1016/j.ifacol.2019.12.499>
- 40
41 469 He K, Gkioxari G, Dollár P, Girshick R (2018) Mask R-CNN. ArXiv170306870 Cs
- 42
43 470 Hothorn T, Bretz F, Westfall P, et al (2023) multcomp: Simultaneous Inference in General Parametric Models
- 44
45 471 Jaccard P (1901) Etude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vaudoise*
46 472 *Sci Nat* 37:547–579
- 47
48 473 Jarvis DE, Ho YS, Lightfoot DJ, et al (2017) The genome of *Chenopodium quinoa*. *Nature* 542:307–312.
49 474 <https://doi.org/10.1038/nature21370>
- 50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 475 Jia W, Tian Y, Luo R, et al (2020) Detection and segmentation of overlapped fruits based on optimized mask R-CNN
4 476 application in apple harvesting robot. Comput Electron Agric 172:105380.
5 477 <https://doi.org/10.1016/j.compag.2020.105380>
- 6 478 Jung A (2022) imgaug: Image augmentation library for deep neural networks
7
8 479 Kang F, Li J, Wang C, Wang F (2023) A Lightweight Neural Network-Based Method for Identifying Early-Blight
9 480 and Late-Blight Leaves of Potato. Appl Sci 13:1487. <https://doi.org/10.3390/app13031487>
- 10 481 Kienbaum L, Correa Abondano M, Blas R, Schmid K (2021) DeepCob: precise and high-throughput analysis of maize
11 482 cob geometry using deep learning with an application in genebank phenomics. Plant Methods 17:91.
12 483 <https://doi.org/10.1186/s13007-021-00787-6>
- 13 484 Landis J (2022) ggsid: Side Grammar Graphics
14
15 485 Lee H-S, Shin B-S (2020) Potato Detection and Segmentation Based on Mask R-CNN. J Biosyst Eng 45:233–238.
16 486 <https://doi.org/10.1007/s42853-020-00063-w>
- 17 487 Lenth RV, Bolker B, Buerkner P, et al (2023) emmeans: Estimated Marginal Means, aka Least-Squares Means
18
19 488 Li Y, Feng X, Liu Y, Han X (2021) Apple quality identification and classification by image processing based on
20 489 convolutional neural networks. Sci Rep 11:16618. <https://doi.org/10.1038/s41598-021-96103-2>
- 21 490 Liu J, Wang X (2020) Tomato Diseases and Pests Detection Based on Improved Yolo V3 Convolutional Neural
22 491 Network. Front Plant Sci 11:
23
24 492 Lozano-Isla F (2021) inti: Tools and Statistical Procedures in Plant Science
25
26 493 Lozano-Isla F, Apaza J-D, Mujica Sanchez A, et al (2023) Enhancing quinoa cultivation in the Andean highlands of
27 494 Peru: a breeding strategy for improved yield and early maturity adaptation to climate change using traditional cultivars.
28 495 Euphytica 219:26. <https://doi.org/10.1007/s10681-023-03155-8>
- 29 496 Machefer M, Lemarchand F, Bonnefond V, et al (2020) Mask R-CNN Refitting Strategy for Plant Counting and Sizing
30 497 in UAV Imagery. Remote Sens 12:3015. <https://doi.org/10.3390/rs12183015>
- 31
32 498 Maharjan S, Alsadoon A, Prasad PWC, et al (2020) A novel enhanced softmax loss function for brain tumour detection
33 499 using deep learning. J Neurosci Methods 330:108520. <https://doi.org/10.1016/j.jneumeth.2019.108520>
- 34
35 500 Maliro MFA, Guwela VF, Nyaika J, Murphy KM (2017) Preliminary Studies of the Performance of Quinoa
36 (*Chenopodium quinoa* Willd.) Genotypes under Irrigated and Rainfed Conditions of Central Malawi. Front Plant Sci
37 501 8:. <https://doi.org/10.3389/fpls.2017.00227>
- 38
39 503 Patiranage DS, Rey E, Emrani N, et al (2022) Genome-wide association study in quinoa reveals selection pattern
40 504 typical for crops with a short breeding history. eLife 11:e66873. <https://doi.org/10.7554/eLife.66873>
- 41
42 505 R Core Team (2020) R: A language and environment for statistical computing. Vienna, Austria
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

- 1
2
3 506 Ramamurthy K, Thekkath RD, Batra S, Chattopadhyay S (2023) A novel deep learning architecture for disease
4 classification in Arabica coffee plants. *Concurr Comput Pract Exp* 35:e7625. <https://doi.org/10.1002/cpe.7625>
- 5 507
6 508 Revelle W (2021) psych: Procedures for Psychological, Psychometric, and Personality Research
7
8 509 Sabouri H, Sajadi SJ, Jafarzadeh MR, et al (2021) Image processing and prediction of leaf area in cereals: A
9 comparison of artificial neural networks, an adaptive neuro-fuzzy inference system, and regression methods. *Crop Sci*
10 510 61:1013–1029. <https://doi.org/10.1002/csc2.20373>
- 11 511
12
13 512 Santis GD, Ronga D, Caradonia F, et al (2018) Evaluation of two groups of quinoa (*Chenopodium quinoa* Willd.)
14 accessions with different seed colours for adaptation to the Mediterranean environment. *Crop Pasture Sci* 69:1264–
15 513
16 514 1275. <https://doi.org/10.1071/CP18143>
- 17
18 515 Schmidt P, Hartung J, Rath J, Piepho H-P (2019) Estimating Broad-Sense Heritability with Unbalanced Data from
19 Agricultural Cultivar Trials. *Crop Sci* 59:525–536. <https://doi.org/10.2135/cropsci2018.06.0376>
- 20
21 517 Schneider CA, Rasband WS, Eliceiri KW (2012) NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*
22 518 9:671–675. <https://doi.org/10.1038/nmeth.2089>
- 23
24 519 Simonyan K, Zisserman A (2015) Very Deep Convolutional Networks for Large-Scale Image Recognition
- 25
26 520 Sosa-Zuniga V, Brito V, Fuentes F, Steinfort U (2017) Phenological growth stages of quinoa (*Chenopodium quinoa*)
27 based on the BBCH scale. *Ann Appl Biol* 171:117–124. <https://doi.org/10.1111/aab.12358>
- 28
29 522 Stanschewski CS, Rey E, Fiene G, et al (2021) Quinoa Phenotyping Methodologies: An International Consensus.
30 Plants 10:1759. <https://doi.org/10.3390/plants10091759>
- 31
32 523
33 524 Tan M, Le QV (2020) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks
- 34
35 525 Tapia M, Gandarillas H, Alandia S, et al (1979) La quinua y la kañiwa: cultivos andinos. Centro Internacional de
36 Investigaciones para el Desarrollo (CIID), Instituto Interamericano de Ciencias Agricolas (IICA), Bogotá
- 37
38 527 Wang H (2020) Garbage Recognition and Classification System Based on Convolutional Neural Network VGG16.
39 In: 2020 3rd International Conference on Advanced Electronic Materials, Computers and Software Engineering
40 (AEMCSE). pp 252–255
- 41
42 529
43 530 Warman C, Fowler JE (2021) Deep learning-based high-throughput phenotyping can drive future discoveries in plant
44 reproductive biology. *Plant Reprod* 34:81–89. <https://doi.org/10.1007/s00497-021-00407-2>
- 45
46 532 Wickham H, Chang W, Henry L, et al (2023) ggplot2: Create Elegant Data Visualisations Using the Grammar of
47 Graphics
48 533
49
50 534 Wrigley CW, Corke H, Seetharaman K, Faubion J (2015) Encyclopedia of Food Grains. Academic Press
- 51
52 535 Xie X, Ma Y, Liu B, et al (2020) A Deep-Learning-Based Real-Time Detector for Grape Leaf Diseases Using
53 Improved Convolutional Neural Networks. *Front Plant Sci* 11:
54 536
55
56
57
58
59
60

- 1
2
3 537 Yu D, Zha Y, Sun Z, et al (2023) Deep convolutional neural networks for estimating maize above-ground biomass
4 538 using multi-source UAV images: a comparison with traditional machine learning algorithms. *Precis Agric* 24:92–113.
5 539 <https://doi.org/10.1007/s11119-022-09932-0>
6
7 540 Zhou Y, Srinivasan S, Mirnezami SV, et al (2019) Semiautomated Feature Extraction from RGB Images for Sorghum
8 541 Panicle Architecture GWAS. *Plant Physiol* 179:24–37. <https://doi.org/10.1104/pp.18.00974>
9
10 542 Zystro J, Colley M, Dawson J (2018) Alternative Experimental Designs for Plant Breeding. In: *Plant Breeding*
11 543 Reviews. John Wiley & Sons, Ltd, pp 87–117
12
13 544
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

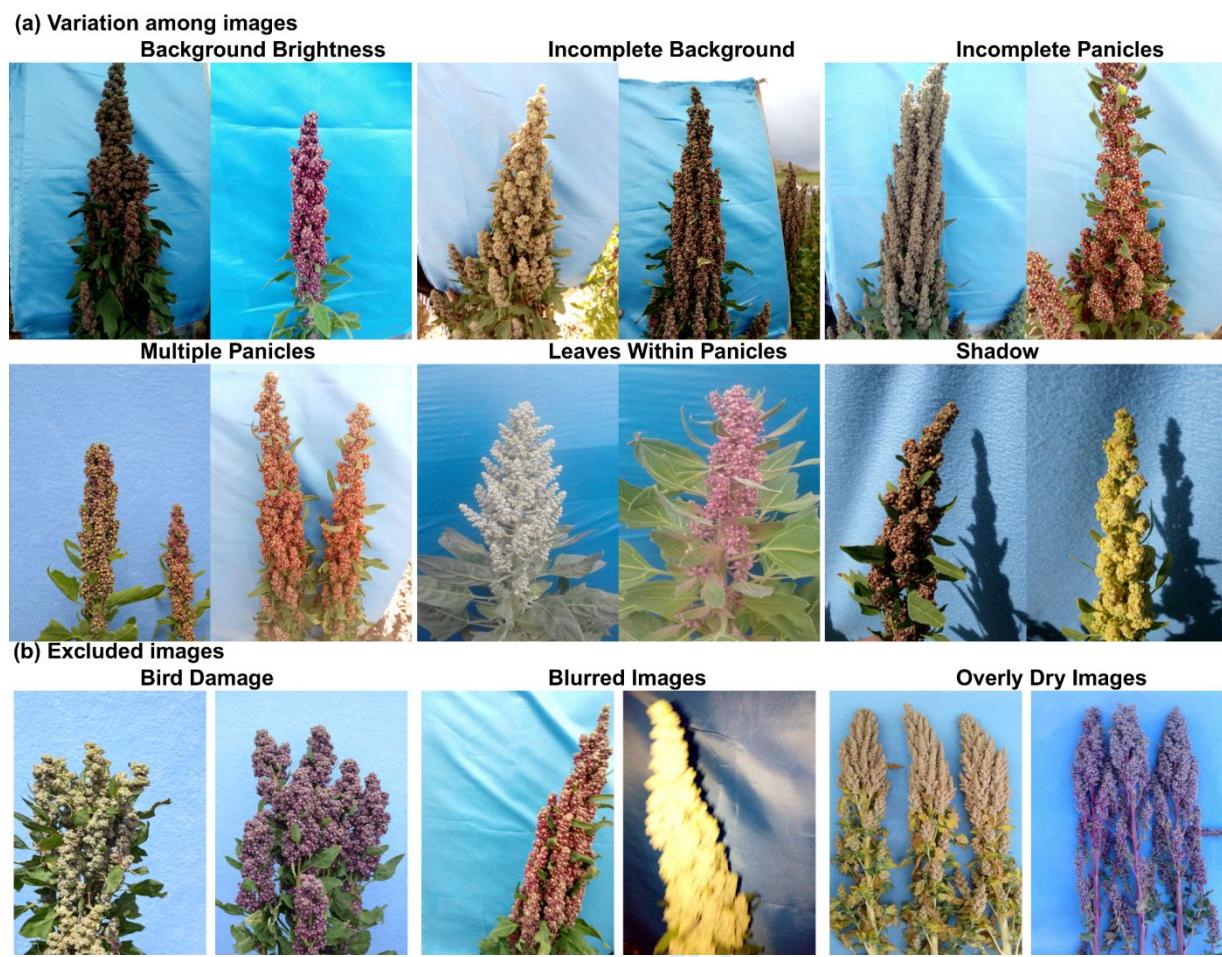


Fig. 1: Different types of heterogeneity in the image dataset of quinoa panicles. (a) Types of variation included in the analysis: Variation of lightning conditions, background, and panicle frame in the pictures. (b) Types of variation excluded in the analysis: Bird-damaged, blurred images, and overdry panicles were excluded from the dataset to avoid bias during training and image analysis.

545

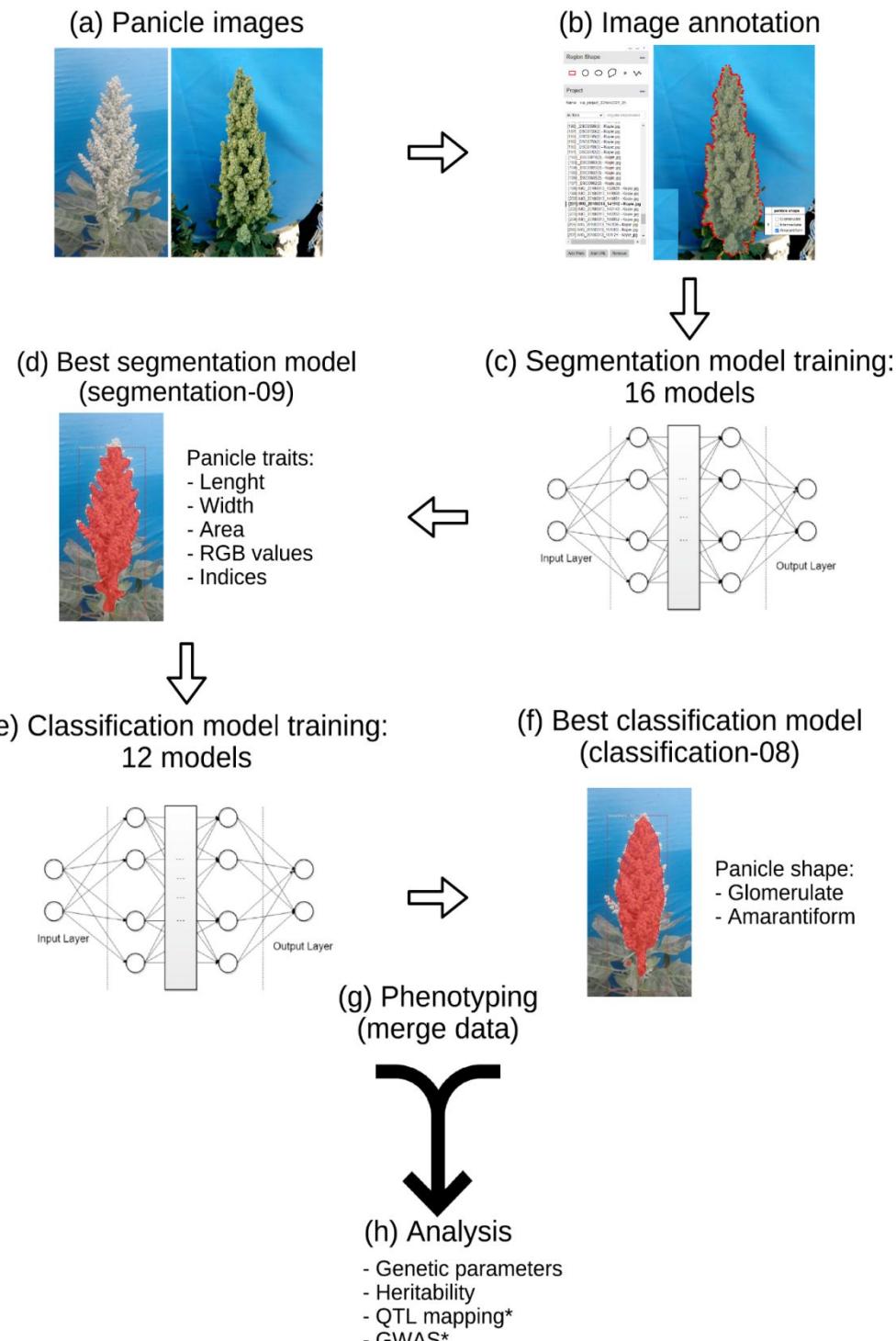


Fig. 2: Quinoa panicle image analysis pipeline using deep learning for segmentation and classification (a) Pictures from the field experiments (b) Image annotation using via software for segmentation model training (c) Model training for image segmentation (d) Best model for segmentation and trait extraction (e) Model training for image classification (f) Best model for classification (g) Merge of phenotypic data extracted from images (h) Data analysis and further application. The asterisk denotes analyses that were not conducted in the present study.

546

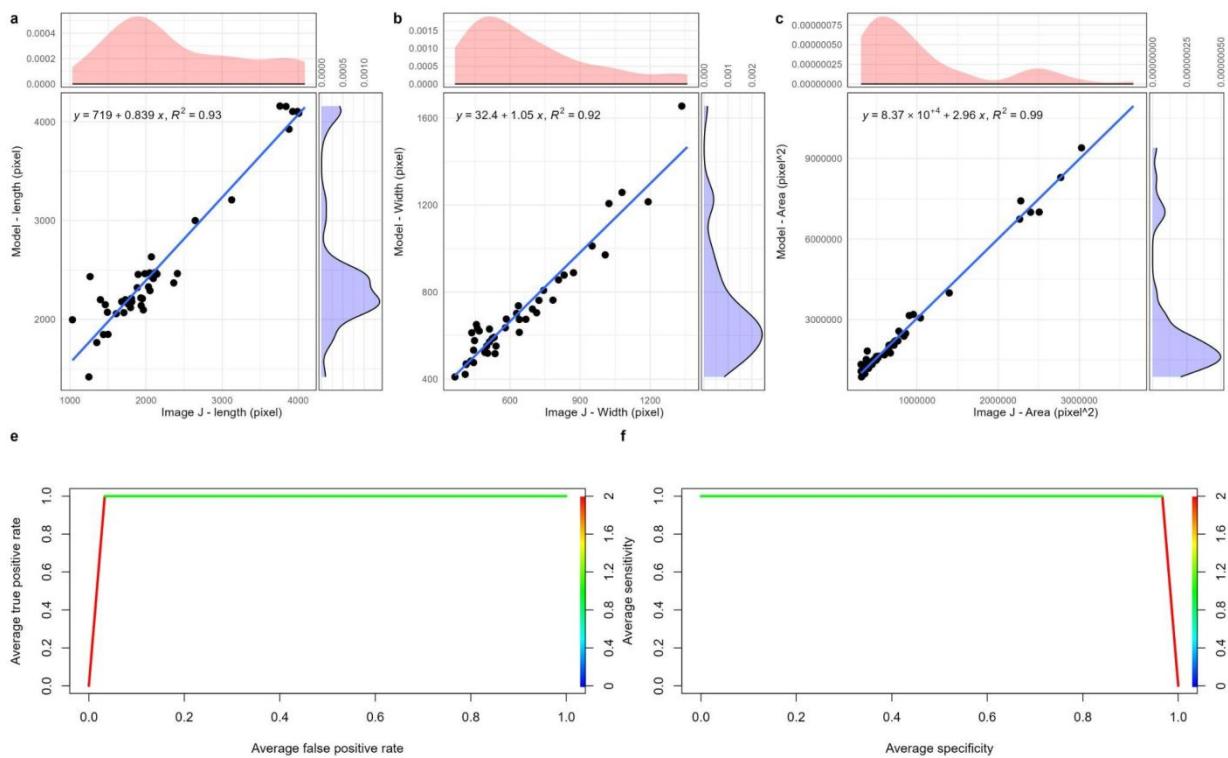


Fig. 3: Comparison between ImageJ and deep learning pipeline. (a-b) Regression analysis for panicle length, width, and area. (e-f) ROC curve for average true and false positive rate. ROC curve for sensitivity and specificity. Fifteen images were taken from the selection and registration trials ($n = 60$).

547

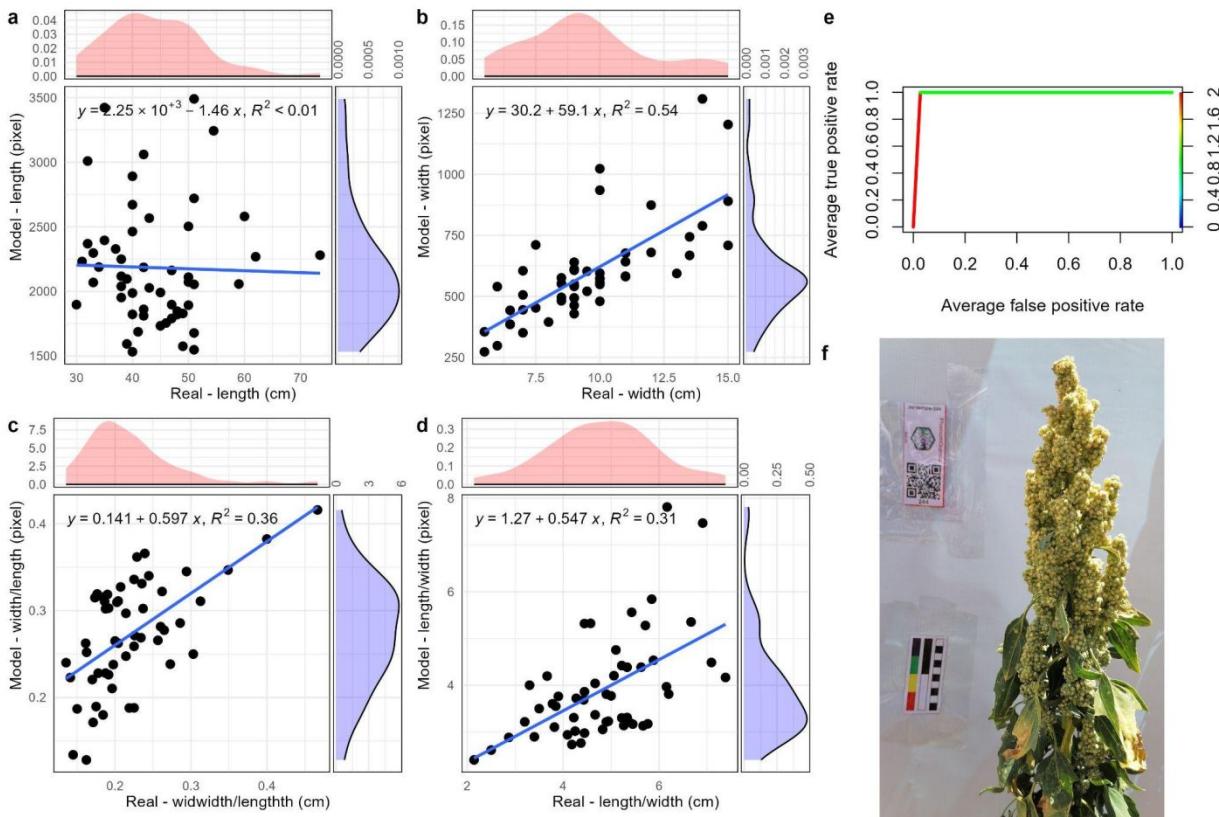


Fig. 4: Model prediction ability with images with scale under field conditions. (a-d) Regression analysis for panicle traits and indices. (e) ROC curve for sensitivity and specificity. (f) Panicle picture under field conditions with scale and QR code. The pictures were taken during the seed production trials ($n = 106$).

548

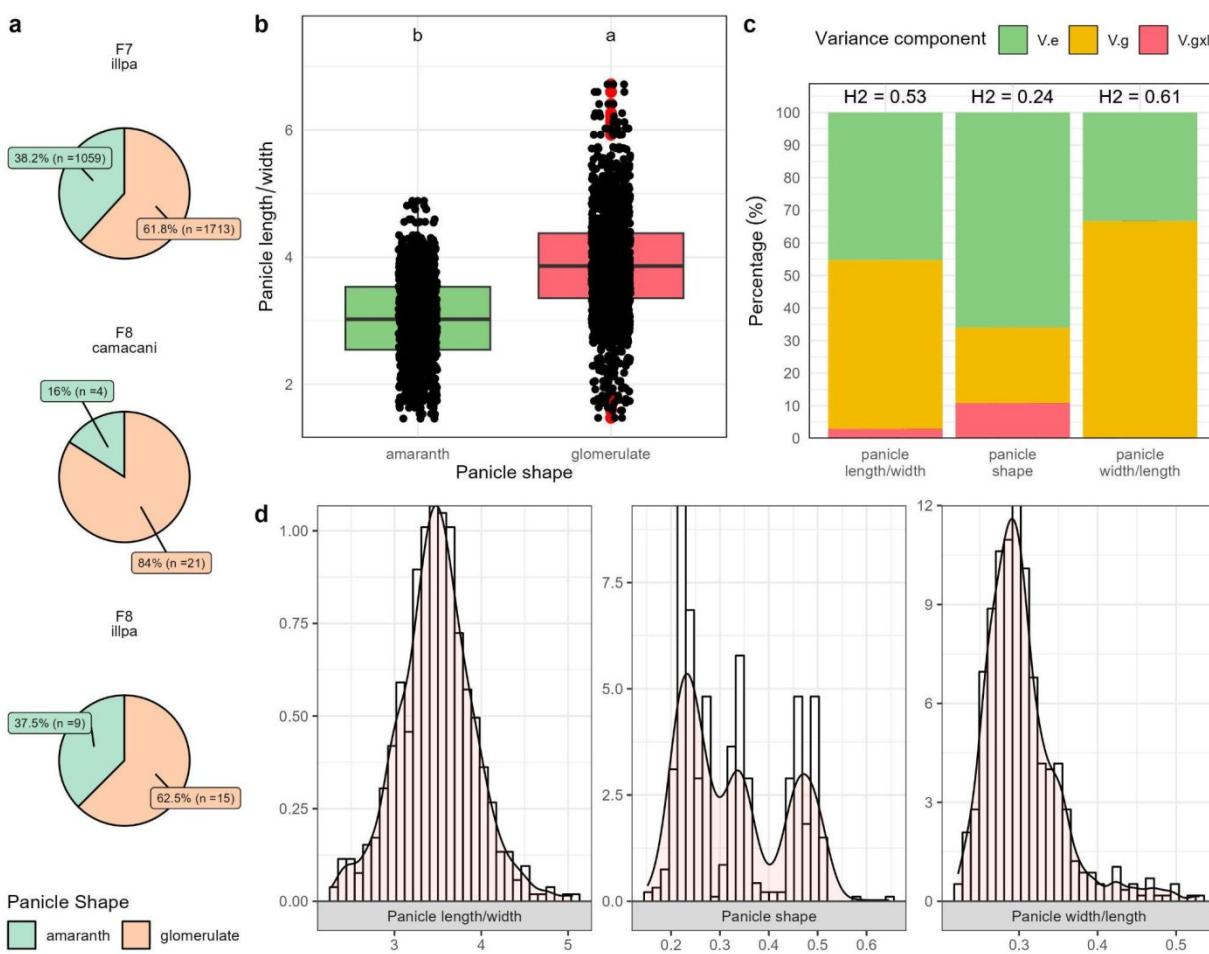


Fig. 5: Quantitative-genetic analysis from quinoa panicle traits under multi-location trials. (a) Panicle shape distribution across generation and location. (b) Panicle relation between the shape and the ratio length and width. (c) Variance partition and heritability from panicle traits in quinoa. (d) Trait distribution after second stage analysis for panicle shape and indices. Values are represented by the BLUPs ($n = 548$).

549

1
2
3 550 Table 1: Images from multi-environmental trials from 2016 to 2019 employed for image analysis using high-throughput phenotyping for quinoa panicles.
4 551 Experiments were conducted under different experimental designs (i.e. lattice and randomized complete blocks - RCBD) with presence of unbalanced data. Devices
5 552 and image resolution were different for each experiment. The images from seasons 2021-2022 were used for model validation with the use of a scale in the images.
6
7 553

Trial	Season	Genotypes	Generation	Location	Exp. design	Device	Resolution	Pictures	Scale
Selection	2016-2017	1200	F6	Camacani, Puno	RCBD	Nikon D5101	2736 x 3648	3862	no
Registration	2017-2018	600	F7	Illpa, Puno	Lattice 10x10	ZTE Blade A610	2448 x 3264	1240	no
Registration	2018-2019	25	F8	Camacani, Puno	Lattice 5x5	Samsung SM-T285M	1440 x 2560	25	no
Registration	2018-2019	25	F8	Illpa, Puno	Lattice 5x5	Samsung SM-T285M	1440 x 2560	25	no
Production	2021-2022	57	F10	Camacani, Puno	RCBD	Motorola G30	3456 x 4608	108	yes

1

2

3
556
557
558

Table 2: Model performance for 16 segmentation models using Mask R-CNN for extracting the phenotypic information using different mask resolution, loss weight, and neural network parameters. Values represented by Least-squares means (ap595) and standard error (ste). Significance (sig) was estimated based on Tukey ($p < 0.05$). Each model was run with five replications.

Model	loss weight	mask resolution	heads.m	ap595	ste	sig
segmentation-1	mask1class10	28x28	all	0.763	0.012	a
segmentation-2	mask20class1	28x28	all	0.754	0.012	a
segmentation-3	mask10class1	28x28	all	0.747	0.012	a
segmentation-4	mask1class1	28x28	all	0.738	0.012	a
segmentation-5	mask1class1	28x28	heads	0.734	0.012	a
segmentation-6	mask20class1	28x28	heads	0.727	0.012	ab
segmentation-7	mask10class1	28x28	heads	0.725	0.012	ab
segmentation-8	mask1class10	28x28	heads	0.686	0.012	b
segmentation-9	mask10class1	56x56	all	0.801	0.012	a
segmentation-10	mask20class1	56x56	all	0.8	0.012	a
segmentation-11	mask1class1	56x56	all	0.791	0.012	a
segmentation-12	mask1class10	56x56	all	0.782	0.012	a
segmentation-13	mask1class1	56x56	heads	0.758	0.012	a
segmentation-14	mask1class10	56x56	heads	0.751	0.012	a
segmentation-15	mask10class1	56x56	heads	0.737	0.012	ab
segmentation-16	mask20class1	56x56	heads	0.704	0.012	b

36

559

38

560

40

41

42

43

44

45

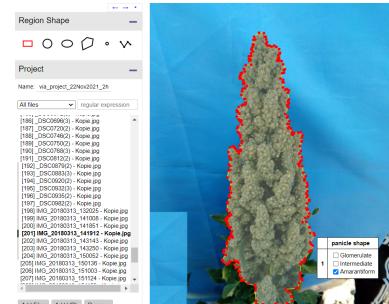
46

47

1
2
3 561 Table 3: Model performance for 12 classification models using three different architectures, dense layers, and activation function configurations for binary image
4 562 classification for quinoa panicles in amaranthiform and glomerulate panicle shapes. Values represented by Least-squares means (accuracy) and standard error (ste).
5 563 Significance (sig) was estimated based on Tukey ($p < 0.05$). Each model was run with five replications.
6
7

Model	architecture	dense layers	activation function	accuracy	ste	sig
classification-01	InceptionV3	128	sigmoid	0.917	0.013	a
classification-02	InceptionV3	128	softmax	0.871	0.013	a
classification-03	InceptionV3	1024	sigmoid	0.909	0.015	a
classification-04	InceptionV3	1024	softmax	0.924	0.015	a
classification-05	VGG16	128	sigmoid	0.909	0.015	a
classification-06	VGG16	128	softmax	0.909	0.015	a
classification-07	VGG16	1024	sigmoid	0.901	0.019	a
classification-08	VGG16	1024	softmax	0.909	0.015	a
classification-09	EfficientNetB0	128	sigmoid	0.909	0.015	a
classification-10	EfficientNetB0	128	softmax	0.901	0.019	a
classification-11	EfficientNetB0	1024	sigmoid	0.901	0.019	a
classification-12	EfficientNetB0	1024	softmax	0.854	0.015	a

(a) Panicle images



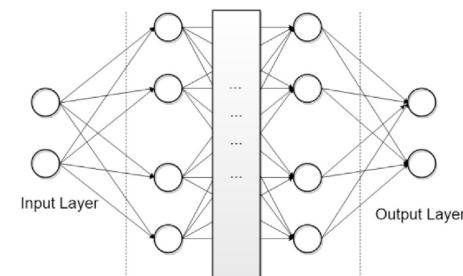
(d) Best segmentation model
(segmentation-09)



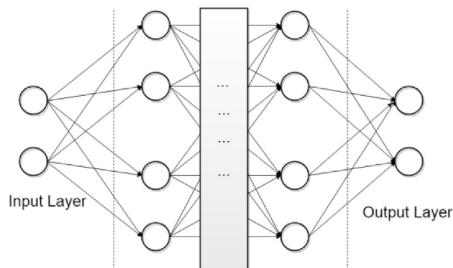
- Panicle traits:
 - Length
 - Width
 - Area
 - RGB values
 - Indices



(c) Segmentation model training:
16 models



(e) Classification model training:
12 models

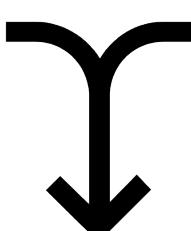


(f) Best classification model
(classification-08)



Panicle shape:
- Glomerulate
- Amarantiform

**(g) Phenotyping
(merge data)**



(h) Analysis

- Genetic parameters
 - Heritability
 - QTL mapping*
 - GWAS*

Supplementary Information

A high-throughput phenotyping pipeline for quinoa (*Chenopodium quinoa*) panicles using image analysis with convolutional neural networks

Flavio Lozano-Isla¹, Lydia Kienbaum¹, Bettina I.G. Haussmann^{1,2}, Karl Schmid^{1*}

Supplementary Figure 1

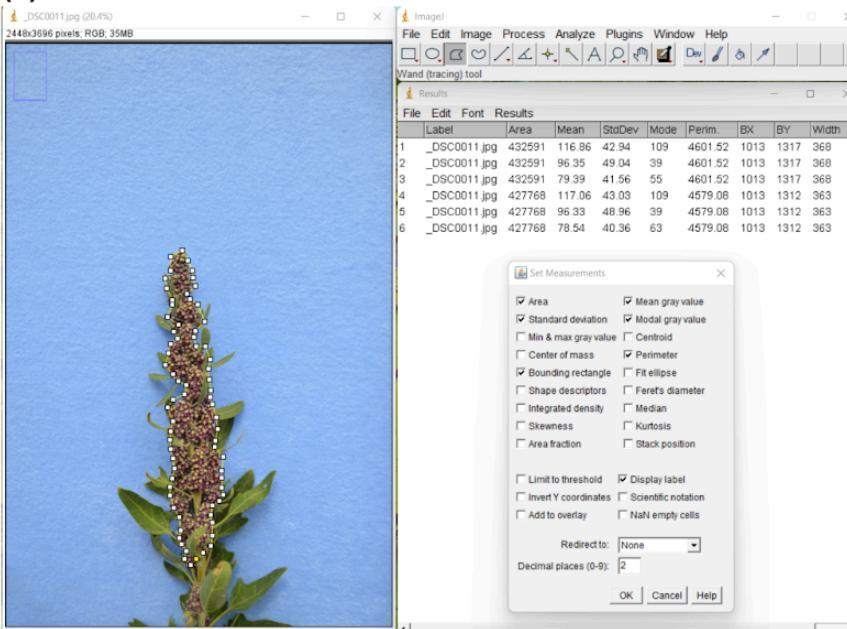
Figure S1: Picture extraction for each experimental unit with blue background under field conditions.

Supplementary Figure 2

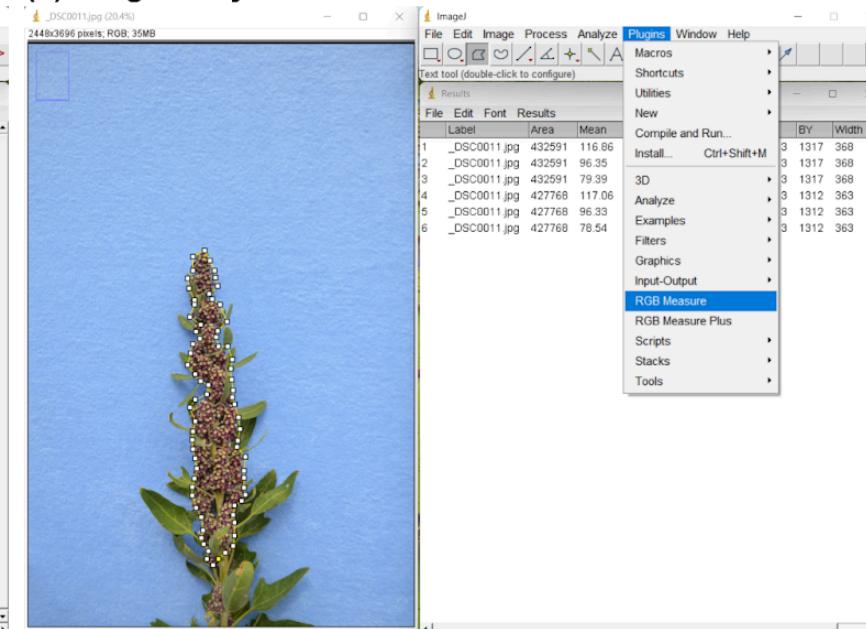
Figure S2: Panicle image analysis. (a) Image from field condition with scale. (b) Image annotation is based on seven classes: panicle, scale, label, barcode, and each channel from RGB spectra

1
2
3
4 **Supplementary Figure 3**
5
6
7
8

(a) Parameters selection



(b) Image analysis



31 **Figure S2:** Manual feature extraction using ImageJ. (a) Selection of parameters, width, length, and area. (b) To extract RGB channel values, the RGB measure
32 plugin was used.
33
34
35
36
37
38
39
40
41
42
43
44
45
46

1 **Journal**
2 **Title**
3
4
5 **STable1**
6 **STable2**
7 **STable3**
8 **STable4**
9 **STable5**
10 **STable6**
11 **STable7**
12 **STable8**
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

1 Plant Breeding
2 A high-throughput phenotyping pipeline for quinoa (*Chenopodium quinoa*) panicles using image analysis with
3
4 Analysis of variance for segmentation model
5 Model segmentation parameters and replications
6 Analysis of variance for classification model
7 Model classification parameters and replications
8 Model selection for stagewise analysis to calculate the BLUPs in the second stage. Formula notation for fixed
9 First stage analysis for variance components for F7 IIIpa season with replication by plots
10 Quantitative-genetic analysis from quinoa panicle traits under multi-location trials. Variance partition and
11 Model comparison using a linear mixed model (lmer) and a binomial generalized linear mixed model (glmer) for
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

For Peer Review

	Factor	Sum Sq	Df	F value	Pr(>F)	Sig
3	(Intercept)	44.99094195	1	60193.5877	6.83535E-97	***
4	loss	0.001387589	3	0.618820454	0.605360542	ns
5	heads.m	0.038802305	1	51.91378135	8.11413E-10	***
6	mask_resolution	0.019764416	1	26.44290297	2.77492E-06	***
7	loss:heads.m	0.005134886	3	2.289995943	0.086716287	ns
8	loss:mask_resolution	0.002791439	3	1.24489302	0.300863554	ns
9	heads.m:mask_resolution	0.002887349	1	3.862997635	0.053704616	ns
10	loss:heads.m:mask_resolution	0.009011006	3	4.018622085	0.011043601	*
11	Residuals	0.047835997	64			

14

15 Significance:

16 0.001 **** 0.01 *** 0.05 **

For Peer Review

	model	loss	heads.m	mask_resolution	ap595
1	segmentation-16mask20class1 all		56x56		0.832
2	segmentation-16mask20class1 all		56x56		0.815
3	segmentation-10mask1class1 all		56x56		0.814
4	segmentation-12mask10class1 all		56x56		0.812
5	segmentation-11mask1class10 all		56x56		0.809
6	segmentation-12mask10class1 all		56x56		0.804
7	segmentation-12mask10class1 all		56x56		0.801
8	segmentation-12mask10class1 all		56x56		0.799
9	segmentation-10mask1class1 all		56x56		0.798
10	segmentation-16mask20class1 all		56x56		0.795
11	segmentation-11mask1class10 all		56x56		0.792
12	segmentation-12mask10class1 all		56x56		0.790
13	segmentation-11mask1class10 all		56x56		0.788
14	segmentation-8 mask1class10 all		28x28		0.788
15	segmentation-16mask20class1 all		56x56		0.787
16	segmentation-10mask1class1 all		56x56		0.785
17	segmentation-10mask1class1 all		56x56		0.783
18	segmentation-11mask1class10 all		56x56		0.781
19	segmentation-9 mask10class1 all		28x28		0.781
20	segmentation-10mask1class1 all		56x56		0.778
21	segmentation-4 mask1class1 heads		56x56		0.772
22	segmentation-16mask20class1 all		56x56		0.772
23	segmentation-15mask20class1 all		28x28		0.772
24	segmentation-5 mask1class10 heads		56x56		0.771
25	segmentation-15mask20class1 all		28x28		0.767
26	segmentation-8 mask1class10 all		28x28		0.765
27	segmentation-8 mask1class10 all		28x28		0.762
28	segmentation-4 mask1class1 heads		56x56		0.760
29	segmentation-4 mask1class1 heads		56x56		0.759
30	segmentation-1 mask1class1 heads		28x28		0.757
31	segmentation-7 mask1class1 all		28x28		0.756
32	segmentation-7 mask1class1 all		28x28		0.756
33	segmentation-5 mask1class10 heads		56x56		0.755
34	segmentation-5 mask1class10 heads		56x56		0.755
35	segmentation-13mask20class1 heads		28x28		0.755
36	segmentation-15mask20class1 all		28x28		0.755
37	segmentation-6 mask10class1 heads		56x56		0.753
38	segmentation-8 mask1class10 all		28x28		0.753
39	segmentation-14mask20class1 heads		56x56		0.753
40	segmentation-9 mask10class1 all		28x28		0.751
41	segmentation-4 mask1class1 heads		56x56		0.750
42	segmentation-9 mask10class1 all		28x28		0.750
43	segmentation-6 mask10class1 heads		56x56		0.749
44	segmentation-9 mask10class1 all		28x28		0.749
45	segmentation-4 mask1class1 heads		56x56		0.748
46	segmentation-2 mask1class10 heads		28x28		0.745
47	segmentation-8 mask1class10 all		28x28		0.744

1	segmentation-5 mask1class10 heads	56x56	0.740
2	segmentation-13mask20class1heads	28x28	0.740
3	segmentation-15mask20class1 all	28x28	0.738
4	segmentation-11mask1class10 all	56x56	0.738
5	segmentation-1 mask1class1 heads	28x28	0.738
6	segmentation-14mask20class1heads	56x56	0.737
7	segmentation-3 mask10class1heads	28x28	0.737
8	segmentation-15mask20class1 all	28x28	0.737
9	segmentation-3 mask10class1heads	28x28	0.736
10	segmentation-5 mask1class10 heads	56x56	0.735
11	segmentation-13mask20class1heads	28x28	0.734
12	segmentation-2 mask1class10 heads	28x28	0.734
13	segmentation-7 mask1class1 all	28x28	0.732
14	segmentation-3 mask10class1heads	28x28	0.732
15	segmentation-6 mask10class1heads	56x56	0.732
16	segmentation-1 mask1class1 heads	28x28	0.731
17	segmentation-13mask20class1heads	28x28	0.730
18	segmentation-6 mask10class1heads	56x56	0.729
19	segmentation-7 mask1class1 all	28x28	0.726
20	segmentation-1 mask1class1 heads	28x28	0.726
21	segmentation-6 mask10class1heads	56x56	0.722
22	segmentation-2 mask1class10 heads	28x28	0.721
23	segmentation-1 mask1class1 heads	28x28	0.720
24	segmentation-3 mask10class1heads	28x28	0.719
25	segmentation-7 mask1class1 all	28x28	0.719
26	segmentation-14mask20class1heads	56x56	0.714
27	segmentation-9 mask10class1all	28x28	0.704
28	segmentation-3 mask10class1heads	28x28	0.701
29	segmentation-14mask20class1heads	56x56	0.680
30	segmentation-13mask20class1heads	28x28	0.677
31	segmentation-2 mask1class10 heads	28x28	0.654
32	segmentation-14mask20class1heads	56x56	0.637
33	segmentation-2 mask1class10 heads	28x28	0.578
34			
35			
36			
37			
38			
39			
40			
41			
42			
43			
44			
45			
46			
47			
48			
49			
50			
51			
52			
53			
54			
55			
56			
57			
58			
59			
60			

	Factor	Sum Sq	Df	F value	Pr(>F)	Sig
3	(Intercept)	11.04054032	1	15323.7045	5.46201E-34	***
4	architecture	0.001216088	2	0.843934215	0.442903162	ns
5	`dense layers`	7.18898E-05	1	0.099779371	0.754943186	ns
6	`activation function`	0.001220879	1	1.694517312	0.205894278	ns
7	architecture:`dense layers`	0.003658977	2	2.539236359	0.10082659	ns
8	architecture:`activation function`	0.00239744	2	1.663762008	0.211424305	ns
9	`dense layers`:`activation function`	0.000178134	1	0.247240744	0.623744759	ns
10	architecture:`dense layers`:`activation functic	0.003661559	2	2.541028037	0.100678732	ns
11	Residuals	0.016571217	23			

14 ---

15 Significance:

16 0.001 **** 0.01 *** 0.05 **

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

32

33

34

35

36

37

38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

For Peer Review

model	architecture	num epoch	dense layers	activation function	accuracy	
classification-09	EfficientNetB0 1	200	128	sigmoid	0.894	
classification-09	EfficientNetB0 1	200	128	sigmoid	0.909	
classification-09	EfficientNetB0 1	200	128	sigmoid	0.924	
classification-10	EfficientNetB0 2	200	128	softmax	0.894	
classification-10	EfficientNetB0 2	200	128	softmax	0.909	
classification-11	EfficientNetB0 3	200	1024	sigmoid	0.894	
classification-11	EfficientNetB0 3	200	1024	sigmoid	0.909	
classification-12	EfficientNetB0 4	200	1024	softmax	0.818	
classification-12	EfficientNetB0 4	200	1024	softmax	0.833	
classification-12	EfficientNetB0 4	200	1024	softmax	0.909	
classification-01	InceptionV3	1	200	128	sigmoid	0.894
classification-01	InceptionV3	1	200	128	sigmoid	0.909
classification-01	InceptionV3	1	200	128	sigmoid	0.924
classification-01	InceptionV3	1	200	128	sigmoid	0.939
classification-02	InceptionV3	2	200	128	softmax	0.833
classification-02	InceptionV3	2	200	128	softmax	0.849
classification-02	InceptionV3	2	200	128	softmax	0.894
classification-02	InceptionV3	2	200	128	softmax	0.909
classification-03	InceptionV3	3	200	1024	sigmoid	0.894
classification-03	InceptionV3	3	200	1024	sigmoid	0.909
classification-03	InceptionV3	3	200	1024	sigmoid	0.924
classification-04	InceptionV3	4	200	1024	softmax	0.909
classification-04	InceptionV3	4	200	1024	softmax	0.924
classification-04	InceptionV3	4	200	1024	softmax	0.939
classification-05	VGG16	1	200	128	sigmoid	0.894
classification-05	VGG16	1	200	128	sigmoid	0.909
classification-05	VGG16	1	200	128	sigmoid	0.924
classification-06	VGG16	2	200	128	softmax	0.894
classification-06	VGG16	2	200	128	softmax	0.909
classification-06	VGG16	2	200	128	softmax	0.924
classification-06	VGG16	2	200	128	softmax	0.924
classification-07	VGG16	3	200	1024	sigmoid	0.894
classification-07	VGG16	3	200	1024	sigmoid	0.909
classification-08	VGG16	4	200	1024	softmax	0.864
classification-08	VGG16	4	200	1024	softmax	0.909
classification-08	VGG16	4	200	1024	softmax	0.955

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

```
1  
2      Model df      AIC  
3      1  NA  NA  
4      2  7       1395.04  
5      3  NA  NA  
6      4  4       1412.48  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
```

For Peer Review

Fixed model

trait ~ 0 + (1|location) + (1|year) + (1|line:location) + (1|line:year) + (1|line:year:location) + (1|line)
trait ~ 0 + (1|location) + (1|year) + (1|line:location) + (1|line:year) + (1|line)
trait ~ 0 + (1|trial) + (1|line:trial) + (1|line)
trait ~ 0 + (1|trial) + (1|line)

For Peer Review

```
1  
2     Modelo Formula  
3     lmer    panicle_shape ~ 1 + (1|trial) + (1|line:trial) + (1|line)  
4     glmer  panicle_shape ~ 1 + (1|trial) + (1|line:trial) + (1|line), family = binomial(link = "lo  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60
```

For Peer Review

	V.g	V.gxI	V.e	H2.s
3	0.078086	0.0638616	0.11118	0.662292
4	7.129707	9.609889	NA	0.689994

For Peer Review

	Trait	geno	env	year	mean	std	min	max	V.g	V.e	repeatability
3	panicle_length.width	547	1	1	3.49	0.84	1.159	6.53	0.6	0.26	0.822
4	panicle_width.length	545	1	1	0.31	0.08	0.17	0.68	0.01	0	0.856
5	panicle_shape	548	1	1	0.45	0.44	-0.13	1.23	0.16	0.1	0.756

For Peer Review

1	Trait	geno	env	year	mean	std	min	max	V.g	V.gxl	V.e	H2.s	H2.c
2	panicle_length.width	547	2	2	3.52	0.83	1.193	6.57	0.36	0.02	0.31	0.88	0.53
3	panicle_width.length	542	2	2	0.3	0.08	0.164	0.68	0	0	0	0.91	0.61
4	panicle_shape	548	2	2	0.35	0.44	-0.24	1.12	0.05	0.02	0.13	0.63	0.24

For Peer Review