



HAL
open science

An automatic non-invasive classification for plant phenotyping by MRI images: An application for quality control on cauliflower at primary meristem stage

Yifan Zhou, Raphaël Maître, Mélanie Hupel, Gwenn Trotoux, Damien Penguilly, François Mariette, Lydia Bousset, Anne-Marie Chèvre, Nicolas Parisey

► To cite this version:

Yifan Zhou, Raphaël Maître, Mélanie Hupel, Gwenn Trotoux, Damien Penguilly, et al.. An automatic non-invasive classification for plant phenotyping by MRI images: An application for quality control on cauliflower at primary meristem stage. *Computers and Electronics in Agriculture*, 2021, 187, pp.106303. 10.1016/j.compag.2021.106303 . hal-03347927

HAL Id: hal-03347927

<https://hal.inrae.fr/hal-03347927>

Submitted on 2 Aug 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

1 **Manuscript number** COMPAG_2020_2352_R2

2

3 With: 2 Tables; 8 Figures (online only color); 1 Appendix

4

5 **Title** An automatic non-invasive classification for plant phenotyping by MRI images:
6 an application for quality control on cauliflower at primary meristem stage

7

8 Yifan Zhou^a, Raphaël Maître^d, Mélanie Hupel^b, Gwenn Trotoux^a, Damien Penguilly^c, François
9 Mariette^b, Lydia Bousset^a, Anne-Marie Chevre^a and Nicolas Parisey^a,

10

11 *^aInstitute for Genetics Environment and Plant Protection (IGEPP), Institut National de*
12 *recherche pour l'agriculture, l'alimentation et l'environnement (INRAE), Domaine de la*
13 *Motte, 35653, Le Rheu, France*

14 *^bINRAE OPAALE, 17 avenue de Cucillé, CS 64427- 35044 Rennes, France*

15 *^cCATE - Station expérimentale, Vezendoquet, 29250 Saint Pol de Léon, France*

16 *^dOBS - Organisation Bretonne de Sélection, Kernonen, 29250 Plougoum, France*

17

18 **Corresponding Author:** Nicolas Parisey UMR1349IGEPP, INRAE Domaine de la motte

19 35 653 Le Rheu cedex, France Tél 02 23 48 51 95 Fax : 02 23 48 51 50

20 nicolas.parisey@inrae.fr

21

22 **Highlights**

23 * An automatic non-invasive method detects cauliflower curd deformation

24 * Tomographic images analysed by machine learning and deep learning methods

25 * Depending on the plant developmental stages, cross-validated F1-score were up to 95%

26 * On combined developmental stages, cross-validated F1-score is 88.67 %.

27

28 **Abstract**

29 During the past few years, milder autumn and winter seasons have caused severe problems to
30 cauliflower harvest of Brittany region in France, mainly due to curd deformation.

31 Consequently, cauliflower breeders are working on breeding new varieties that are more
32 robust to climate change to stabilize the quality of cauliflower production. The aim of this
33 study was to identify at which stage of the curd formation, significant difference can be
34 detected between healthy and stressed cauliflower. A non-invasive classification based on
35 Magnetic Resonance Imaging (MRI) images for cauliflower phenotyping was proposed. Plants
36 exposed to vernalization stress were sampled at different times around primary meristem stage,
37 then both MRI imaged and apex dissected. A work flow was developed to extract features
38 from MRI images. A classification on phenotype was learned by LDA, QDA, PLSDA and
39 CNN binary classification between two groups: healthy and stressed cauliflower. Promising
40 F1 score and MCC up to 95% were achieved. Curd deformation is the main cause for
41 cauliflower's later physiological disorders when reaching maturity. Therefore, the cauliflowers
42 with deformation could be removed at the earliest, e.g., screening for plant breeding. At the
43 same time, the healthy cauliflowers are not destroyed and continue their life cycle.

44

45 **Keywords** plant phenotyping ; non-invasive classification ; cauliflower primary meristem ;
46 MRI application ; discriminant analysis

47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72

Abstract

During the past few years, milder autumn and winter seasons have caused severe problems to cauliflower harvest of Brittany region in France, mainly due to curd deformation. Consequently, cauliflower breeders are working on breeding new varieties which are more robust to climate change to stabilize the quality of cauliflower production. The aim of this study was to identify at which stage of the curd formation, significant difference can be detected between healthy and stressed cauliflower. A non-invasive classification based on Magnetic Resonance Imaging (MRI) images for cauliflower phenotyping was proposed. Plants exposed to vernalization stress were sampled at different times around primary meristem stage, then both MRI imaged and apex dissected. A work flow was developed to extract features from MRI images. A classification on phenotype was learned by LDA, QDA,PLSDA.

73 Statistical analysis was then applied for a binary classification between two groups: healthy
74 and stressed cauliflower. Promising F1 score and MCC up to 95% were achieved. Curd
75 deformation is the main cause for cauliflower's later physiological disorders when reaching
76 maturity. Therefore, the cauliflowers with deformation could be removed at the earliest, e.g.,
77 screening for plant breeding. At the same time, the healthy cauliflowers are not destroyed and
78 continue their life cycle.

79

80 **Keywords:** plant phenotyping ; non-invasive classification ; cauliflower primary meristem ;
81 MRI application ; discriminant analysis

82

83 **1 Introduction**

84 According to the Food and Agriculture Organization of the United Nations, large-scale
85 experiments in crop phenotyping are a key factor in meeting the future agricultural needs to
86 feed the world and provide biomass for energy while using less water and fertilizer under a
87 constantly evolving environment adapted to climate change (Minervini et al., 2015). However,
88 current assessments of phenotypic characteristics for disease resistance or stress in breeding
89 programs rely largely on visual scoring by experts, which is laborious and dull, not sufficiently
90 objective or destructive (Busemeyer et al., 2013). Various imaging methodologies are being
91 used to collect data for quantitative studies of complex traits related to growth, yield and
92 adaption to biotic or abiotic stress (disease, insects, drought and salinity) (Li et al., 2014).
93 There is an urgent need to develop reliable computer vision methods that can extract
94 phenotypic information from experiments at scales from single cell to whole plant, in the
95 greenhouse or on the field (Li et al., 2014). The extracted information, integrated with genetic
96 and environmental data by novel models based on accurate, robust and automatic statistical

97 analysis will give new opportunity to genetic diversity screening, new breeding strategies in
98 agriculture as well as market management.

99 Given the rapid development of high-throughput genotype screening in plant breeding and
100 genomics for related growth, yield and tolerance to different biotic and abiotic stresses, there
101 is a call for more effective and reliable phenotyping data to support modern genetic crop
102 improvement (Li et al., 2014). To accomplish this goal, more and more projects for plant
103 phenotype unite expertise from biological science, computer science, mathematics and
104 engineering. Such an approach was needed to offer cauliflower breeders efficient screening
105 methods on plant development as early as possible, by associating plant phenotype with
106 genomes in imaging systems of computer vision.

107 In recent years, cauliflower winter harvest for the region of Brittany in France has been
108 observed to be very unstable and extremely reduced due to warmer autumns (Tremellat, 2017).
109 Physiological disorders, such as open, ricey (Watts, 1966) or bracty (Kop et al., 2003) head
110 appear. A healthy head is tightly compact with only florets and forms one bracts (Fig. 1a),
111 whereas an open head has gaps among florets (Fig. 1b); a ricey head has protruding flower
112 buds (Fig. 1c); and a bracty head has leaves intermingled with florets (Fig. 1d). These
113 deformations renders cauliflower heads unmarketable, resulting in important commercial
114 losses, e.g. with about a third of harvest was unmarketable (Tremellat, 2017).

115 To induce flowering and thus curd formation, beyond the juvenile phase, cauliflower must
116 be exposed to vernalization at "relative cold" temperature (Wurrand Fellows, 2000). In the
117 couple of weeks following vernalization, cauliflower primary meristem undergoes curd
118 formation period, divided into 4 stages (Kieffer et al., 1998): vegetative, curd-induction, curd-
119 forming and curd-thickening stage, noted as Stage 1, 2, 3 and 4, respectively in this paper.
120 This curd formation period is critical for cauliflower growth. If deformation appears during
121 this period, the cauliflower head will remained deformed during subsequent growth until

122 maturity, about 2 months later. Breeding cauliflowers less sensible to autumn temperature
123 fluctuation is thus desirable to stabilize yields in autumn to winter harvests. To render such
124 breeding possible, early stage phenotyping during curd formation is needed.

125 Floral initiation in cauliflower is the result of fine regulation of a whole network of genes
126 and regulatory loops with interplay between transcription factors (Goslin et al., 2017). This
127 regulation interacts with vernalization (Matschegewski et al., 2015). Bracting in cauliflower
128 depends on its genotype (Kop et al., 2003) and the climate during floral initiation. The effect
129 of temperature and developmental stage on bracting and riciness quality defects have already
130 been studied in the field, either during harvest time (Grevsen et al., 2003) or by a destructive
131 sampling with scanning electron microscopy (Fujime and Okuda, 1996).

132 Using non-invasive methods, healthy cauliflower without deformation could be kept for
133 further growth. However, at this moment, the apex of cauliflower meristem is only around
134 0.5mm of diameter, still tightly wrapped in a bunch of huge leaves making it invisible to
135 naked eyes (Fig. 2) and preventing the use common RGB cameras to capture high resolution
136 pictures on meristem without destroying the plant. For organs inavailable from the outside
137 as plant apex, among the available techniques one could either use external imaging e.g.
138 spectrometry or hyperspectral imaging; or internal imaging e.g. Xray or MRI. External
139 imaging would be possible if the external part of the plant has features correlated to the apex
140 deformation. At such early stages as the one investigated in this study, experts are not able
141 to assess the healthy state by external observation of plant morphology. Raman
142 spectrometry has been investigated in a companion project but will not be discussed here as
143 no success was achieved. Current internal imaging techniques are Xray and Magnetic
144 Resonance Imaging (MRI). Xray relies on a difference in tissue density, whereas MRI is
145 based on the relaxation time of the tissues. Specifically, in plants the structure and chemical
146 composition of the tissues influences the water molecules relaxation time (Musse and Van-

147 As, 2018). Even though Xrays are cheaper and easier to use, this makes MRI a more
148 selective technique to detect differences in tissues structure among plants. Although in most
149 applications MRI is used to investigate water relations and transport in plant tissues, MRI
150 can also be used to measure other plant constituents, such as metabolites and air spaces.
151 MRI therefore provides access to a wide range of information about plant tissues, including
152 structural characteristics at different length scales and physical-chemical features (Musse
153 and Van-As, 2018). We hypothesised that such changes might occur in the transition from
154 vegetative to floral induction in plant apex. Recent advances in spectroscopic techniques,
155 such as Magnetic Resonance Imaging (MRI) promise non-invasive measurements on plant
156 (Rascher et al., 2013) allowing assessment of plant phenotype revealing plant's inner part
157 without destroying its outside part. However, due to apex's tiny size, its MRI image has
158 only few pixels, which prevents an efficient image analysis. As the stems with deformed
159 meristem have a form different from that of healthy one (Hupel, 2018), this offers the
160 prospect of phenotyping by image analysis of stem shapes.

161 The aim of this study was to identify at which stage of the curd formation, significant
162 difference can be detected between healthy and stressed cauliflower. Plants exposed to
163 vernalization stress were sampled at different times around primary meristem stage, then both
164 MRI imaged and apex dissected. A work flow was developed to extract features from MRI
165 images. Statistical analysis was then applied for a binary classification between two groups:
166 healthy and stressed cauliflower. Since this application is designed to satisfy industrial need,
167 focused was set on how to solve practical problems encountered during different steps in work
168 flow, at same time, proposing adequate and efficient models.

169 **2 Materials and methods**

170 **2.1 Data collection**

171 Two environment conditions were imposed in greenhouse to simulate a normal autumn
172 (group H) and a stressful warmer autumn weather (group S). The latter one was to simulate a
173 milder autumn leading to a shorter vernalization period for cauliflower. A F1 hybrid variety of
174 cauliflower detected as sensitive to temperature fluctuations when cultivated at large-scale
175 Brittany region of France was selected for the experiment. Cauliflower seeds were sown in
176 June 2018 on a mixture of sand and vermiculite at 20°C, with plants in group S sown two
177 weeks later than plants in group H. Their seedlings at two-leaf stage were transferred to a
178 plastic greenhouse at a temperature of 12-13°C for plant hardening during 7 weeks. The plants
179 were transplanted in 7.5L pots and and grown for 7 further weeks, then vernalized at 4°C for
180 either 2 weeks (group S) or 4 weeks (group H).

181 After the vernalization, plants in group H were cultivated under tunnels with an average
182 daily temperature between 9 and 15°C preventing the risk of devernalization, whereas those in
183 group S were kept in a greenhouse with a temperature between 15 and 20°C simulating a
184 warmer autumn. In this way, it was expected that the seeds of the same homogeneous variety
185 would grow to healthy cauliflower heads under normal autumn, called group H, but those
186 under stressful autumn would grow to stressed ones, called group S, in this paper.

187 The plants were sent sampled for MRI measurement weekly between 0 and 31 days after
188 the end of vernalization, 5 plants per date and per treatment each time, in order to have
189 samples distributed on the 4 stages of primary meristem.

190 The MRI measurement was carried out by a 1.5 Tesla MRI whole body scanner
191 (Magnetom Avanto, Siemens, Erlangen, Germany) equipped with an eight-channel "knee"
192 receiver coil. The plant was laid down on the examination table because it was too high to be
193 placed uprightly (Fig. 3). Due to this reason, cauliflower apex might not be found in the coil

194 center due to gravity. This was especially true for plant of meristem at stage 1 and 2 due to its
195 young and fragile stem. Another practical issue was that plant leaves were piled up in the ring
196 resulting in aliasing artifacts in MRI. Therefore, some extra image pre-processing methods
197 were carried out and will be explained in Section 2.2.1.

198 The 3D MRI images were acquired by a 3D turbo spin echo (TSE3D) sequence with a voxel
199 size of $0.5\text{mm}\times 0.5\text{mm}\times 0.5\text{mm}$, a matrix size of 192×192 , a FOV of 96×96 , 96 slices per
200 volume, a slice thickness of 0.5mm, an echo time of 9.5ms, a repetition time of 500ms, 2
201 averages, a turbo factor of 14 and a bandwidth of 263 Hz/pixel. These values were chosen as
202 the best compromise between on the one hand enhancing, at best, the contrast between the
203 stem and the rest of the plant by the MRI operator at the time of acquisition and on the other
204 hand, the acquisition time short enough to allow analysis of a sufficient number of plants per
205 day.

206 The acquisition time for one plant was about 33 min. Plant were well watered at the eve of
207 every acquisition in order to improve MRI image contrast.

208 After the MRI acquisition, cauliflower was dissected and photographed to enable breeding
209 experts to identify growth stage and thus construct a ground truth for the database. Dissected
210 cauliflower apex were stained with aceto-carmin, observed under a magnifying lens (Nikon
211 SMZ-U, zoom 1:10) and digital RGB images were taken (Fig. 4). In this article, four stages
212 were distinguished: 1. Vegetative stage; 2. Curd-induction stage; 3. curd-forming stage; 4.
213 Curd-thickening stage. The top line (images ABCD) in Figure 4 illustrates a schematic
214 representation of cauliflower floral induction based on the scanning electron micrographs in
215 (Kieffer et al., 1998). At vegetative stage, only leaf scales are produced (Fig. 4A). During the
216 curd-induction stage, an enlarging empty area becomes visible between leaf scales (Fig. 4B).
217 Through the curd-forming stage, round floral primordia appears at axil of each bract scale (Fig.
218 4C). The growth of bracts scales is repressed at curd-thickening stage (Fig. 4D). For further

219 growth, the apex consists of only floral primordia stopped in their development (Smyth1995),
220 grown into florets (ramified group of flowers). Their further maturation into flowering is
221 postponed long afterwards.

222 The middle (images EFGH) and bottom (images IJKL) line in Figure 4 give corresponding
223 examples of each of the stages on healthy and stressed cauliflower, respectively. Altogether,
224 100 plants were collected for the year 2018 and 60 plants for 2019.

225 These RGB images were used for a double-blinded identification of meristem
226 developmental stage, compared manually with Fig. 4 by experts on cauliflower's meristem
227 morphological development in order to decide their corresponding sample's primary meristem
228 stage.

229

230 **2.2 Feature extraction**

231 For each dissected plant, called a sample in this paper, a set of MRI images was acquired,
232 called raw images. Feature extraction consisted of three steps: extraction of region of interest,
233 image of contour and image of skeleton.

234 **2.2.1 Extraction of region of interest**

235 Its aim was to select slice from raw images (Fig. 5) and pre-process the selected slice. There
236 were 96 raw images on plane XY per sample. The raw image resolution was 192×192 pixels.
237 Therefore, a sample can be represented in 3D (Fig. 5b) with the apex of cauliflower meristem
238 circled in red. In this 3D presentation, the top of a cauliflower plant has of volume of 0.5dm³.
239 Reminding that an apex was only about 0.5mm, it barely corresponded to one pixel on raw
240 image. Hence, it was chosen to extract features on cauliflower's stem appearance.

241 Because variable numbers of leaves had been included in the coil during MRI acquisition,
242 rather than extracting 3D morphological features directly from raw images, the plant 3D

243 morphological information was contained within the 2D images on plane XZ (Fig5c) and YZ
244 around the stem apex (Fig5d). Hence, the database consisted of 320 slices for 160 samples.

245 If the plant apex was perfectly in coil center during MRI acquisition, the two middle slices
246 on plane XZ and YZ were the best illustration. However, when the plant apex was not
247 perfectly in coil center, the cauliflower meristem was missed or occluded in middle slice (Fig.
248 6a,b) Therefore, a manual selection was required in order to find one best slice on each plane
249 which illustrated the meristem as clearly and entirely as possible (Figure 6c,d).

250 Aliasing artifacts of MRI acquisition happened when one leaf exceeding scanner's field of
251 view was partially projected onto the other side of image, (Fig. 6e). Aliasing is an artifact that
252 can occur in MRI images acquisition when the scanned object is larger than the square image
253 area that is to be measured, which is called the field of view (FOV). As a consequence of
254 sampling issues, portions of the object outside of the desired FOV get mapped to an incorrect
255 location inside the FOV. For our 2D slices, this artefact reproject leaves with much higher
256 contrast than the rest of the stem. In order to work properly on the stem, beside the presence of
257 this artefact, we improve the local contrast and enhance the definitions of edges everywhere in
258 the slice by using histogram equalization. Hence, a contrast limited adaptive histogram
259 equalization (CLAHE) that enhance contrast by using information in the vicinity of each pixel
260 while putting limitations (i.e. maxima) to the extent of said contrast augmentation (Gonzalez
261 and Woods, 2008) was applied on this kind of slices (Fig. 6f).

262

263 To focus on cauliflower stem apex morphology, and remove unnecessary plant leaves and
264 petioles, an extraction of region of interest (ROI) was performed as follows Fig. 7a to 7c). A
265 contour extraction method based on Otsu's thresholding (Otsu, 1979; Gonzalez and Woods,
266 2006) was firstly used to find object of interest, generating an image called mask in this paper
267 (Fig. 7b). The slice on the mask was scaled to a higher resolution, from 192×96 to 384×192

268 pixels. To improve contrast inside ROI, the Contrast Limited Adaptive Histogram
269 Equalization algorithm (CLAHE) (Pizer et al., 1987) was carried out on scaled image. This
270 final ROI image (Fig. 7c), was ready for further feature extraction.

271 **2.2.2 Image of contour**

272 To extract morphological features on cauliflower stem apex, the same contour method
273 (Gonzalez and Woods, 2006) was again applied on the ROI image. Only the contour with the
274 largest area was considered as the final object of interest (Fig. 7d), called image of contour in
275 this paper. Five categories of features were calculated: contour marked in green, rectangle in
276 blue, hull, ellipse in red and intensity (Fig. 7d). For each category, 3 or 5 features were
277 computed (Table 1), chosen in a way that the features' value was invariant to object position in
278 the image.

279 **2.2.3 Image of skeleton**

280 To extract morphological features on cauliflower main stem, an image of skeleton was
281 produced as follows. From a ROI image (Fig. 8a), the corresponding mask (Fig. 8b) was first
282 morphologically thinned with a maximum iteration of 10 by (Zhang and Suen, 1984), (Fig. 8c).
283 It was then skeletonized to 1 pixel (Fig. 8d). The difference between the thinned and
284 skeletonized image was the image of skeleton (Fig. 8e), having a similar form to plant main
285 stem. Three categories of features were calculated on image of skeleton: contour marked in
286 green, rectangle in blue and hull (Table 1).

287 **2.3 Binary classification with discriminant analysis**

288 The classification issue was binary, with the two groups of cauliflower either healthy or
289 stressed. The database consisted of 160 plants, with 320 images, distributed on 4 stages (Table
290 2, see columns Stage, Size, H, S) with nearly half of plants in stage 4,. Due to this limited data
291 size, a **leave-one-out cross-validation** (Devijver and Kittler, 1982) was applied to assess the
292 predictive capability of the classifiers.

293 Several supervised algorithms were tested for learning and validation steps based on
294 features extracted from image of contour or/and image of skeleton: Linear Discriminant
295 Analysis (LDA) (Fisher, 1936), Quadratic Discriminant Analysis (QDA) (Hastie et al., 2009)
296 and Partial Least Squares Discriminant Analysis (PLSDA) (Barker and Rayens, 2003). Simple
297 classifiers were chosen because of covariates multicollinearity, small samples and unbalance
298 in our dataset can be troublesome to highly non-linear and/or complex classifiers.

299 The application had significant ratios of feature number to sample size. For example, the
300 number of features for the chosen classifiers on image of contour and skeleton on stage 3 is 34
301 versus 26 slices from 13 samples. In order to avoid overfitting and multicollinearity problems
302 in machine learning (Burnham and Anderson, 2002), LDA was used with automatic
303 shrinkage by Ledoit-Wolf lemma (Ledoit and Wolf, 2004), QDA with regularized covariance
304 (Friedman,1989) and we used PLSDA.

305 The regularization parameter for QDA and the adaptive component number for PLSDA
306 were automatically chosen by a nested leave-one-out cross-validation with the inner layer to
307 find hyper-parameters giving best F1 score on subsamples and the outer layer to evaluate
308 algorithm performance on the whole cross-validated data.

309 F1 score (F1) (Rijsbergen, 1979) Matthews Correlation Coefficient (MCC) (Matthews,
310 1975) and Jaccard Index (JI) (Jaccard, 1912) were used to evaluate the classifiers performance.
311 The MCC was a balanced measure of the quality of binary classification with +1 a perfect
312 prediction, 0 no better than random prediction and -1 a total disagreement between prediction
313 and observation. The JI was defined as the size of the intersection divided by the size of the
314 union of two label sets which are here the predicted set of labels and the observed set of labels.
315 These metrics are adapted to the situation where two classes were of very different size
316 (Chicco and Jurman, 2020) (Table 2, see different size between "H" and "S").

317

318 **2,4 Binary classifications with deep learning**

319 To complement our classifications based on selected features, deep learning was implemented,
320 a class of computational models composed of multiple processing layers learning
321 representations of data with multiple levels of abstraction (LeCun et al., 2015). Those
322 algorithms have proved to be very efficient in a wide variety of domains, most notably
323 computer vision (Emmert-Streib et al., 2020). Among those algorithms, Convolutional Neural
324 Networks (CNNs) are well known for their success in many computer vision tasks such as
325 image classification (Krizhevsky et al., 2012) and objects recognition (Li et al., 2015). Deep
326 learning was used to explore two main questions. First, could a CNN, using slices, lead to
327 higher scoring than our classifications based on selected features on said slices? Second, could
328 a CNN, using directly the 3D volumes, lead to good scoring?

329 To explore the ability of a CNN to outperform other classifiers, on selected slices, transfer
330 learning was used, a technique to re-purpose a previously trained model (Yosinski et al., 2014).
331 A classical Xception architecture (Chollet et al, 2017) was used as base model, pretrained on
332 the large generalist ImageNet dataset (Deng et al., 2009) with more than 14 million images of
333 thousands of categories. The fully connected layers were removed (and associated multiclass
334 problem) and the rest of the convolutional layers were used as fixed feature extractors to feed
335 a new neural network of two fully connected layers of 512 neurons. From here on, this CNN
336 will be referred to as the 2D CNN. After the first transfer-learning step, subsequent fine-tuning
337 i.e. training also the convolutional layers was also tried. In both cases, given our small datasets,
338 on-the-fly data augmentation with rotations and axial symmetries was used.

339 To assess the properties of a CNN using 3D volumes, a readily available architecture
340 previously used on CT Scans for binary classification in human epidemiology (Zunair et al.,
341 2020) was adapted. From here on, this CNN will be referred to as the 3D CNN. Due to the
342 relative scarcity of volumic datasets and their divergence from our use case, transfer learning

343 was not chosen and the 3D CNN was trained from scratch. Again, on-the-fly data
344 augmentation with rotations was used. All deep learnings were performed by dividing our
345 dataset into training (70%) and validation (30%) sets.

346

347 **2.5 Hardware and librairies used**

348 The codes in the application were written in python using library OpenCV (Bradski, 2014),
349 skimage (van der Walt et al., 2014), scikit-learn (Pedregosa et al. 2011) and tensorflow (Abadi
350 et al., 2015). The calculation was carried out on a common desktop Dell Precision Tower 3420
351 with Intel Xeon E3-1225 v6, 8192KB cache and an NVIDIA Tesla K80.

352 The processing time for feature extraction was quite negligible, no more than several
353 millisecond per image. The computation time for supervised learning and validation depended
354 on sample and feature size. For example, it took less than 1 second for a cross-validation by
355 LDA with features of image contour and skeleton (34 features) on stage 234 (232 images for
356 116 samples). Computation time for CNNs were up to an hour for 3D CNN over 200 epochs.
357 For all our application, from pre-processing steps to deep learning, codes and a sample dataset
358 are available upon request to the corresponding author.

359 **3 Results**

360 In order to decide from which stage the classifiers can distinguish healthy from stressed
361 cauliflower apex, the classifiers were firstly learned and validated on samples of the year 2018
362 and 2019 together, but on separate individual stages, 1, 2, 3 or 4 (Table 2 upper lines). On all
363 the individual stages except on stage 1 (when cauliflower meristem was still on vegetative
364 state), one or several classifiers could reach expectation (Boughorbel et al., 2017) with F1
365 above 85% and MCC above 65% (Table 2, marked in italics; see below for the rationale
366 behind our expectation threshold). Therefore, the classifiers could distinguish healthy from
367 stressed cauliflower as early as from curd-induction stage.

368

369 One or 2 classifiers with best performance on every stage are marked with an asterisk. LDA
370 gives most of the best performance (marked with an asterisk) compared to QDA and PLSDA
371 (Table 2). Besides, most of the best performance by LDA are computed by features from
372 image of contour (Table 2). As the computation of the JI do not lead to significantly different
373 conclusions than the ones based on F1 and MCC, in order to keep Table 2 comprehensible, we
374 made JI scores available in Table 1 of the Appendix.

375

376 All the results reaching expectation are marked in italics, with F1 above 85% and MCC above
377 65% (Boughorbel et al., 2017). In the referenced article (Boughorbel et al, 2017), there are
378 classifiers built on many different data sets (Table 5; 46 data sets). For a generalist classifier
379 (in their case, the « SVM.imb » column), the average MCC obtained is 67.76%. We rounded
380 roughly to 65% and considered that a good result would be strictly above this approximated
381 mean score. For the F1-score, under the assumption of a class-balanced dataset, the score for a
382 naïve classifier, that always predict the same class whatever the covariate values, is 66 %. If
383 we consider such score as the least we can do and given the different levels of imbalance in our
384 different data sets, which affect such minimal F1-score, we took a higher threshold, namely at
385 85%.

386

387 Based on the previous observation, classifiers on mixed stage 234 were computed in order
388 to test whether features were sufficiently different to predict cauliflower healthy state without
389 prior knowledge on meristem stage. The result was quite promising, with 88.67% on F1 and
390 67.93% on MCC based on features of contour and skeleton by QDA (Table 2 middle lines).

391

392 Since the classifiers were calculated by supervised algorithms in machine learning, we
393 wanted to know the influence of sample size on classifier performance, by comparing the
394 largest data set to smaller ones. For this account, classifiers were computed only on samples of
395 the year 2018 on stage 4 and 234 (Table 2, botom lines). In this comparison, the sample size is
396 increased by about 12% and 35% compared to samples of year 2018 and 2019, respectively.
397 With more samples, the classifier performance was improved, between 0.3% to 10.3%.
398 Therefore, it was concluded that the sample size did have an impact on the classifiers'
399 performance.

400

401 Concerning deep learning, both the 2D CNN and the 3D CNN were trained to solve our
402 binary classification problem on stage '234' from 2018 and 2019. This was our biggest
403 relevant dataset even if it was still extremely small by deep learning standards.

404 For the transfer learning of the 2D CNN, a small learning rate ($1e-3$) was chosen and had to
405 stop after a few epochs (20) before overfitting appeared as assessed by visual diagnostic of the
406 loss functions between trained and validated datasets. For the subsequent fine-tuning of the 2D
407 CNN, an even smaller learning rate ($1e-5$) was used and more epochs (50).

408 With transfer learning, the performance of the 2D CNN on the validation dataset were an F1-
409 score of 73.19 % and a MCC of 46.39 %. With subsequent fine-tuning, there were small
410 performance improvement with an F1-score of 75.48 % and a MCC of 50.97 %.

411 For the training of the 3D CNN, a low learning rate of $1e-5$ was used for 50 epochs, afterward
412 overfitting appeared. The performances of the 3D CNN on the validation set were an F1-score
413 of 81.46 % and a MCC of 62.38 %.

414

415 **4 Discussion**

416 In this paper, a non-invasive classification for cauliflower phenotyping by MRI images was
417 proposed. It was an application of screening on cauliflower still at primary stage long before
418 its physiological disorders become visible to naked eyes. More specifically, MRI images were
419 firstly acquired on cauliflower plant with its apex diameter of about 0.5mm, on which features
420 by contour or/and skeleton were extracted. These features were then sent for learning by
421 several discriminant analysis, such as LDA, QDA and PLSDA. The healthy state of
422 cauliflower meristem influenced by temperature fluctuation was then predicted. If deformation
423 had already occurred and would be sufficiently marked, dissecting the apex could give access
424 to deformation. However, it could be the case that the molecular processes have already
425 occurred but the deformation is not yet visible on the apex. Contrasting the two sets of
426 temperature enabled us to detect morphological differences prior to the meristem deformation.
427 However, within each of the two temperature groups, there was variation in the development,
428 making it necessary to assess the actual growth stage of each individual plant. As an
429 alternative to discriminant analysis on selected features, deep learning methods were used to
430 generate predictions.

431 The classifiers could distinguish healthy or stressed cauliflower as early as from curd-
432 induction stage. Experiments showed that cauliflower meristem developed very quickly into
433 stage 1, just in few days after the end of the vernalization period. It might even occur during
434 vernalization if the temperature was not cold enough. Therefore in practice, there would not
435 have many cauliflower on stage 1 for the application on screening, usually scheduled several
436 days after vernalization. Hence, the poor performance on stage 1 of the classifiers was
437 negligible in an industrial context.

438

439 Classifiers on mixed stage 234 were computed in order to test whether features were
440 sufficiently different to predict cauliflower healthy state without prior knowledge on meristem
441 stage. The result was quite promising, with 88.67% on F1 and 67.93% on MCC based on
442 features of contour and skeleton by QDA. Mixing curd induction, forming and thickening
443 stage together could make the application even more automatic, since the only need was to
444 decide a starting day for screening by avoiding meristem vegetative period, e.g. 5 days after
445 the end of vernalization. In this manner, even though cauliflower grows differently due to
446 temperature fluctuation, whether its meristem on stage 2, 3 or 4, will not have important
447 impact on the performance of the prediction. Deep learning results were also promising on
448 mixed stage. Although they were not able to outperform classifiers trained on selected features,
449 their performances were very close. Another interesting result is that the deep learning using
450 volumic data has the best performance among the architecture we tried. It showed on our
451 (unusual) use case the versatility of these approaches.

452

453 Classifiers which were not only efficient but also with stable performance were needed.
454 LDA gave most of the best performance (Table 2, marked with an asterisk) compared to QDA
455 and PLSDA. In several of the experiments reported in Table 2, the F1 or MCC score was
456 lower for the classification using image of contour and skeleton as compared to only image of
457 contour. It could be surprising that the performance of the classifier decreases when more
458 information becomes available to learn from. However, all of our classifiers are using
459 mathematical regularizations (in a way or another) which leads, notably, to some features
460 being nullified (their coefficients set to, nearly, zero) if they do not bring new information to
461 solve the problem. This explains, in part, why the number of available features does not
462 always lead to a high improvement in classification. Besides, most of the best performance by
463 LDA are computed by features from image of contour. Therefore, applying classifier using

464 features from image of contour by LDA for the application of screening on cauliflower at
465 primary meristem stage is suggested.

466 Classifier performance was improved with more samples, between 0.3% to 10.3% (Table
467 2). Therefore, it was concluded that the sample size did have an impact on the classifiers'
468 performance. However, it is likely that discriminant classifiers performances would reach a
469 limit lower than that of state-of the art deep learning performances if enough data were
470 available. Given that data scarcity is a bottleneck, a way forward would be to develop transfer
471 learning and/or fine-tuning of a 3D CNN in a future work.

472 The experimental results showed that models had a rather promising performance on F1
473 score and MCC, especially for the one on features of contour by LDA. The classifiers could
474 provide breeders with elements to decide whether to remove those stressed plants before
475 planting. The associated environmental cost from negative by-product of cultivation on
476 unmarketable plants, such as pesticide, soil resource, water and other energy could be saved.

477 The parameters for the classifiers were calculated completely in an automatic way. The
478 only human intervention was during the extraction of region of interest step, that was to
479 manually re-select slices from MRI raw images when plant did not grow straight upside and to
480 add an histogram equalization on slices when MRI aliasing artefacts occurred. In practice, this
481 manual selection largely depended on whether cauliflower's main stem was strong enough to
482 support plant weight against gravity when laying down. It might be related to meristem young
483 stage, like vegetative and curd-induction stage, but not absolutely. For example, 45% slices
484 were re-selected for data in 2018 versus 98% for 2019. Even with the same genetic type, the
485 plant might grow differently between years due to variable environmental conditions. Given
486 the experiment with 3D CNN, if its performances were to be improved, it would be a good
487 solution to remove the few remaining human interventions.

488 This manual selection problem might also be solved by constructing portable MRI system
489 because image acquisition system's position can be adjusted to plant instead of laying down
490 plant to adapt to fixed system. In fact, MRI bulky, costly and complex hardware limitation is
491 the main factor which prevents it becoming a standard research tool in plant phenotype
492 regardless of its non-invasive advantage. Nevertheless, several MRI mobile prototypes have
493 been constructed from laboratories for potential industrial applications, such as measuring
494 dynamic water change in living stems or fruit (Windt and Blümmler, 2015). A relative low-cost
495 wide bore MRI scanners have also been designed and constructed for rapid quality inspections
496 of fruits and vegetables in order for an industrial food quality assurance and control
497 (McCarthy and Zhang, 2012) (Milczarek and McCarthy, 2012). Hence, constructing mobile
498 MRI systems is a potential research direction of future work.

499 Depending on the plant developmental stages, cross-validated F1-score were up to 95% and
500 on combined developmental stages, cross-validated F1-score was 88.67 % (Table 2) and
501 81,46 % for deep learning. Yet only on one sensor (the MRI) was used for this study, another
502 direction for improvement would be combining multi-modal acquisitions from different
503 sensors, such as chlorophyll fluorescence (Rousseau et al., 2013), with ensemble methods
504 (Zhou, 2012).

505

506 **5 Conclusion**

507 We aimed at improving the early detection of cauliflower curd deformation, the main cause for
508 cauliflower's later physiological disorders when reaching maturity. A non-invasive
509 classification based on Magnetic Resonance Imaging (MRI) images for cauliflower
510 phenotyping was proposed, with tomographic images analysed by machine learning and deep
511 learning methods. Promising F1 score and MCC up to 95% were achieved. Therefore, the
512 cauliflowers with deformation could be removed at the earliest, e.g., screening for plant

513 breeding. At the same time, the healthy cauliflowers are not destroyed and continue their life
514 cycle. We consider this work as another proof of the usefulness and potential of tomographic
515 data for non-invasive plant phenotyping

516 **6 Acknowledgements**

517 This work was part of the project 'GWASSICA' co-funded by the Région Bretagne (France)
518 and EU funds from the FEDER Program (Grant EU00430). A special thank to Mona
519 LOSTANLEN from OBS Innovation for her meticulous work in cultivation, transportation and
520 dissection. We thank Pascale Le-Neve and Patricia Nadan from INRAE Le Rheu for their
521 administrative help.

522

523 **7 CRediT author statement**

524 Conceptualization: AMC, NP, LB, FM; Funding acquisition: AMC, NP, FM; Project
525 administration: NP, AMC, RM, MH; Supervision & Validation: NP, AMC, LB; Investigation:
526 RM, MH; Data curation: YZ, RM, MH, GT, DP; Formal analysis, Methodology: YZ, NP;
527 Software: YZ, NP; Visualization, Writing - original draft: YZ; Resources, Writing - review &
528 editing: all authors.

529

530 **8 Declaration of Interest statement**

531 The authors declare that they have no known competing financial interests or personal
532 relationships that could have appeared to influence the work reported in this paper.

533

534 **References**

535 Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A.,
536 Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y.
537 Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mané, D., Monga, R., Moore, S.,
538 Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P.,
539 Vanhoucke, V., Vasudevan, V., Viégas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke,
540 M., Yu, Y., Zheng, X., 2015. TensorFlow: Large-scale machine learning on heterogeneous
541 systems. [arXiv:1603.04467](https://arxiv.org/abs/1603.04467).

542 Barker, M., Rayens, W., 2003. Partial least squares for discrimination. *Journal of*
543 *Chemometrics* 17, 166 – 173.

544 Boughorbel, S., Jarray, F., El-Anbari, M., 2017. Optimal classifier for imbalanced data using
545 matthews correlation coefficient metric. *PLOS ONE* 12, 1–17.

546 Bradski, G., 2000. The OpenCV Library. *Dr. Dobb's Journal of Software Tools* .

547 Burnham, K.P., Anderson, D.R., 2002. Model Selection and Multimodel Inference: A
548 Practical Information-Theoretic Approach. 2 ed., Springer-Verlag, New York.

549 Busemeyer, L., Mentrup, D., Möller, K., Wunder, E., Alheit, K., Hahn, V., Maurer, H.P., Reif,
550 J.C., Würschum, T., Müller, J., Rahe, F., Ruckelshausen, A., 2013. Breedvision - a multi-
551 sensor platform for non-destructive field-based phenotyping in plant breeding. *Sensors* 13,
552 2830–2847.

553 Chicco, D., Jurman, G., 2020. The advantages of the Matthews correlation coefficient (MCC)
554 over F1 score and accuracy in binary classification evaluation. *BMC Genomics* 21, 6.

555 Chollet, F., 2017. Xception: Deep learning with depthwise separable convolutions. In :
556 Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

557 Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., Fei-Fei, L. 2009. Imagenet: A large-scale
558 hierarchical image database. In: 2009 IEEE Conference on Computer Vision and Pattern
559 Recognition. IEEE, pp. 248–255

560 Devijver, P., Kittler, J., 1982. Pattern Recognition: A Statistical Approach. Prentice-Hall,
561 Englewood Cliffs.

562 Emmert-Streib, F., Yang, Z., Feng, H., Tripathi, S., Dehmer, M., 2020. An introductory
563 review of deep learning for prediction models with big data. *Frontiers in Artificial Intelligence*
564 3.

565 Fisher, R., 1936. The use of multiple measurements in taxonomic problems. *Annals of*
566 *Eugenics* 7, 179–188.

567 Friedman, J.H., 1989. Regularized Discriminant Analysis. *Journal of the American Statistical*
568 *Association* 84, 165.

569 Fujime, Y., Okuda, N., 1996. The physiology of flowering in brassicas, especially about
570 cauliflower and broccoli, 247–254.

571 Gonzalez, R.C., Woods, R.E., 2006. Digital Image Processing (3rd Edition). Prentice-Hall,
572 Inc., Upper Saddle River, NJ, USA.

573 González, R.C., Woods, R.E., 2008. Digital image processing, 3rd Edition. Pearson Education.

574 Goslin, K., Zheng, B., Serrano-Mislata, A., Rae, L., Ryan, P.T., Kwaniewska, K., Thomson,
575 B., Ó'Maoléidigh, D.S., Madueño, F., Wellmer, F., Graciet, E., 2017. Transcription factor
576 interplay between leafy and apetala1/cauliflower during floral initiation. *Plant Physiology* 174,
577 1097–1109.

578 Grevsen, K., Olesen, J., Veierskov, B., 2003. The effects of temperature and plant
579 developmental stage on the occurrence of the curd quality defects "bracting" and "riciness" in
580 cauliflower. *Journal of Horticultural Science and Biotechnology* 78, 638–646.

581 Hastie, T., Tibshirani, R., Friedman, J., 2009. The elements of Statistical Learning. Springer-
582 Verlag New York, NY, USA.

583 Hupel, M., Décembre, 2018. Projet GWASSICA - Suivi de l'initiation florale et de la
584 croissance de la pomme de chou-fleur par IRM. Technical Report. IRSTEA Rennes.

585 Jaccard, P., 1912. The Distribution of the flora in the Alpine zone.1. *New Phytologist* 11, 37–
586 50.

587 Kieffer, M., Fuller, M., J. Jellings, A., 1998. Explaining curd and spear geometry in broccoli,
588 cauliflower and 'romanesco': Quantitative variation in activity of primary meristems. *Planta*
589 206, 34–43.

590 Kop, E., Teakle, G., McClenaghan, E., Lynn, J., King, G., 2003. Genetic analysis of the
591 bracting trait in cauliflower and broccoli. *Plant Science* 164, 803–808.

592 Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep
593 convolutional neural networks. In: *Advances in Neural Information Processing Systems*, pp.
594 1097–1105

595 LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *Nature* 521, 436–444.

596 Ledoit, O., Wolf, M., 2004. Honey, I Shrunk the Sample Covariance Matrix. *The Journal of*
597 *Portfolio Management* 30, 110–119.

598 Li, H., Lin, Z., Shen, X., Brandt, J., Hua, G., 2015. A convolutional neural network cascade
599 for face detection. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern*
600 *Recognition*, pp. 5325–5334

601 Li, L., Zhang, Q., Huang, D., 2014. A review of imaging techniques for plant phenotyping.
602 *Sensors (Basel)* 14, 20078–20111.

603 Matschegewski, C., Zetzsche, H., Hasan, Y., Leibeguth, L., Briggs, W., Ordon, F., Uptmoor,
604 R., 2015. Genetic variation of temperature regulated curd induction in cauliflower: elucidation
605 of floral transition by genome-wide association mapping and gene expression analysis.
606 *Frontiers in Plant Science* 6, 720.

607 Matthews, B., 1975. Comparison of the predicted and observed secondary structure of t4
608 phage lysozyme. *Biochimica et Biophysica Acta (BBA) - Protein Structure* 405, 442 – 451.

609 McCarthy, M., Zhang, L., 2012. *Food Quality Assurance and Control*.

610 Milczarek, R., McCarthy, M., 2012. Cheminform abstract: Low-field mr sensors for fruit
611 inspection. *ChemInform* 43.

612 Minervini, M., Scharr, H., Tsafaris, S.A., 2015. Image analysis: The new bottleneck in plant
613 phenotyping [applications corner]. *IEEE Signal*

614 Musse, M., Van As, H., 2018. NMR imaging of air spaces and metabolites in fruit and
615 vegetable. *Modern Magnetic Resonance*, Webb, G.A., Springer International Publishing,
616 Cham, CHE, p. 1765-1779

617 Otsu, N., 1979. A threshold selection method from gray-level histograms.

618 Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M.,
619 Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher,

620 M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine learning in Python. *Journal of*
621 *Machine Learning Research* 12, 2825–2830.

622 Pizer, S.M., Amburn, E.P., Austin, J.D., Cromartie, R., Geselowitz, A., Greer, T., ter Haar
623 Romeny, B., Zimmerman, J.B., Zuiderveld, K., 1987. Adaptive histogram equalization and its
624 variations. *Computer Vision, Graphics, and Image Processing* 39, 355 – 368.

625 *Processing Magazine* 32, 126–131.

626 Rascher, U., Blossfeld, S., Fiorani, F., Jahnke, S., Jansen, M., Kuhn, A., Matsubara, S., L. A.
627 Martin, L., Merchant, A., Metzner, R., Müller-Linow, M., Nagel, K., Pieruschka, R., Pinto, F.,
628 Kuchendorf, C., Temperton, V., Thorpe, M., Dusschoten, D., Van Volkenburgh, E., Schurr, U.,
629 2013. Non-invasive approaches for phenotyping of enhanced performance traits in bean.
630 *Functional Plant Biology* 38, 968–983.

631 Rijsbergen, C.J.V., 1979. *Information Retrieval*. 2nd ed., Butterworth-Heinemann, Newton,
632 MA, USA.

633 Rousseau, C., Belin, E., Bove, E., Rousseau, D., Fabre, F., Berruyer, R., Guillaumès, J.,
634 Manceau, C., Jacques, M.A., Boureau, T., . High throughput quantitative phenotyping of plant
635 resistance using chlorophyll fluorescence image analysis 9, 17.

636 Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch,
637 S., Rueden, C., Saalfeld, S., Schmid, B., Tinevez, J.Y., White, D.J., Hartenstein, V., Eliceiri,
638 K., Tomancak, P., Cardona, A., 2012. Fiji: an open-source platform for biological-image
639 analysis. *Nature Methods* 9, 676–682.

640 Smyth, D.R., 1995. Flower development: Origin of the cauliflower. *Current Biology* 5, 361 –
641 363.

642 Tremellat, V., 2017. Limiter l’impact des aléas climatiques sur chou-fleur via la création
643 variétale et le développement d’outils de phénotypage. Technical Report. OBS Innovation.

644 van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N.,
645 Gouillart, E., Yu, T., the scikit-image contributors, 2014. scikit-image: image processing in
646 Python. PeerJ 2, e453.

647 Watts, L.E., 1966. Investigations on the inheritance and responses to selection of riceyness in
648 early summer cauliflower. Euphytica 15, 90–98.

649 Windt, C.W., Blümner, P., 2015. A portable NMR sensor to measure dynamic changes in the
650 amount of water in living stems or fruit and its potential to measure sap flow. Tree Physiology
651 35, 366–375.

652 Wurr, D., Fellows, J., 2000. Temperature influences on the plant development of different
653 maturity types of cauliflower. Proc. III Int. Symp. on Brassicas , 69–74.

654 Yosinski, J., Clune, J., Bengio, Y., Lipson, H., 2014. How transferable are features in deep
655 neural networks?. In: Advances in Neural Information Processing Systems, pp. 3320–3328.

656 Zhang, T.Y., Suen, C.Y., 1984. A fast parallel algorithm for thinning digital patterns.
657 Commun. ACM 27, 236–239.

658 Zhou, Z.H., 2012. Ensemble Methods: Foundations and Algorithms. 1st ed., Chapman & Hall.

659 Zunair, H., Rahman, A., Mohammed, N., Cohen, J.P., 2020. Uniformizing techniques to
660 process CT scans with 3D CNNs for tuberculosis prediction, In: Rekik, I., Adeli, E., Park,
661 S.H., Valdés Hernández, M. del C. (Eds.), Predictive Intelligence in Medicine, Lecture Notes
662 in Computer Science. Springer International Publishing, Cham, pp. 156–168.

663

664 **Table 1:** List of features extracted from image of contour or image of skeleton and
 665 corresponding line colour used in Figures 7 and 8.

666 Circle diameter stands for diameter of the circle with same area of the category of the line;

667 Center distance stands for the category's center to contour center;

668 Area ratio stands for the category's area to contour area.

669

Features	1	2	3	4	5	Line colour
Contour	area	perimeter		circle diameter		green
Rectangle	area	perimeter	center distance	width to length	area ratio	blue
Hull	area	perimeter	center distance	circle diameter	area ratio	
Ellipse	area	perimeter	center distance	angle orientation	area ratio	red
Intensity	max	mean	min			

670

671

672 **Table 2:** The performance of classifiers based on features from image of contour or/and
673 skeleton by Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA)
674 and Partial Least Squares Discriminant Analysis (PLSDA) evaluated by F1 score and
675 Matthews Correlation Coefficient (MCC). It is calculated separately on samples at stage 1, 2,
676 3, 4 or at mixed stage on 234; on samples of the year 2018 and 2019. All the results reaching
677 expectation are marked in italics, with F1 more than 85% and MCC more than 65%
678 (Boughorbel et al., 2017). One or 2 classifiers with best performance on every stage are
679 marked with an asterisk. The slice distribution between H and S groups on every stage is also
680 listed.

681

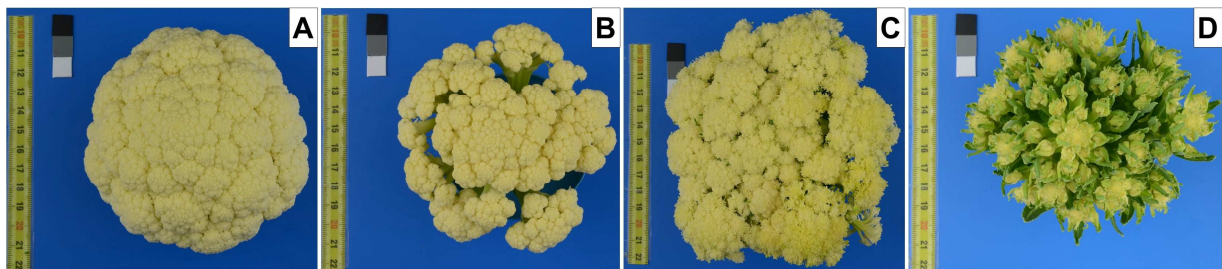
Data	Stage	Features	Size	H	S	LDA F1	LDA MCC	QDA F1	QDA MCC	PLSDA F1	PLSDA MCC
2018 + 2019	1	image of contour	88	22	66	47.37%	34.02%	47.06%	38.24%	84.89%	29.66%
2018 + 2019	1	image of skeleton	88	22	66	57.14%	49.92%	48.15%	27.28%	89.36%*	49.92%*
2018 + 2019	1	image of contour + skeleton	88	22	66	57.14%	43.84%	13.79%	2.42%	86.13%	38.22%
2018 + 2019	2	image of contour	54	34	20	92.75%*	80.00%*	84.85%	61.28%	95.00%*	92.05%*
2018 + 2019	2	image of skeleton	54	34	20	66.67%	0.62%	70.89%	-3.43%	36.84%	2.71%
2018 + 2019	2	image of contour + skeleton	54	34	20	83.78%	50.88%	81.16%	47.88%	73.68%	59.66%
2018 + 2019	3	image of contour	26	20	6	92.68%*	65.92%*	84.44%	-10.95%	54.54%	42.76%
2018 + 2019	3	image of skeleton	26	20	6	80.00%	13.33%	88.37%	37.36%	40.00%	27.24%
2018 + 2019	3	image of contour + skeleton	26	20	6	87.18%	49.08%	86.96%	0.00%	57.14%	42.60%
2018 + 2019	4	image of contour	152	96	56	89.69%*	71.54%*	87.10%	67.04%	74.55%	60.14%
2018 + 2019	4	image of skeleton	152	96	56	89.01%	70.43%	90.26%	72.90%	80.36%	68.90%
2018 + 2019	4	image of contour + skeleton	152	96	56	91.10%*	76.06%*	90.82%	74.29%	85.71%	77.38%
2018 + 2019	234	image of contour	232	150	82	87.30%*	62.64%*	86.01%	64.20%	78.21%	67.41%
2018 + 2019	234	image of skeleton	232	150	82	83.01%	50.22%	86.69%	58.28%	65.41%	47.46%
2018 + 2019	234	image of contour + skeleton	232	150	82	88.03%	64.46%	88.67%*	67.93%*	78.26%	66.74%
2018	4'	image of contour	136	80	56	86.59%	66.37%	83.54%	60.77%	80.00%	66.46%
2018	4'	image of skeleton	136	80	56	88.75%	72.68%	90.68%*	77.18%*	82.88%	71.09%
2018	4'	image of contour + skeleton	136	80	56	88.34%	71.00%	86.39%	64.87%	90.57%*	84.94%*
2018	234'	image of contour	172	94	78	84.38%*	64.73%*	79.10%	57.59%	77.92%	60.05%
2018	234'	image of skeleton	172	94	78	76.76%	49.76%	78.67%	47.72%	70.30%	43.32%
2018	234'	image of contour + skeleton	172	94	78	81.68%	58.85%	83.67%	62.42%	82.05%	67.16%

682

683

684 **Figure 1:** Examples of a healthy head versus stressed cauliflower heads with physiological
685 disorders. The images were captured after they reach a diameter of around 10cm. A. Healthy
686 head is tightly compact with only florets and forms one bracts; B. Open head has gaps among
687 florets; C. Ricey head has protruding flower buds; D. Bracty head has leaves intermingled
688 with florets.

689



690

691

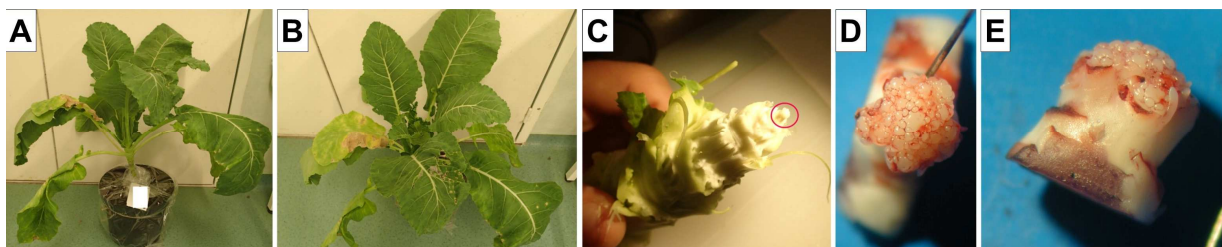
692

693

694

695 **Figure 2:** Example of a cauliflower plant at its primary mersitem stage. A,B. The whole plant
696 has big leaves tightly wrapped around the apex; C. Once leaves are removed, the cauliflower
697 apex becomes visible (circled in red); D,E. On the apex, floral primordia can be examined,
698 here stained with carmine red for better contrast.

699

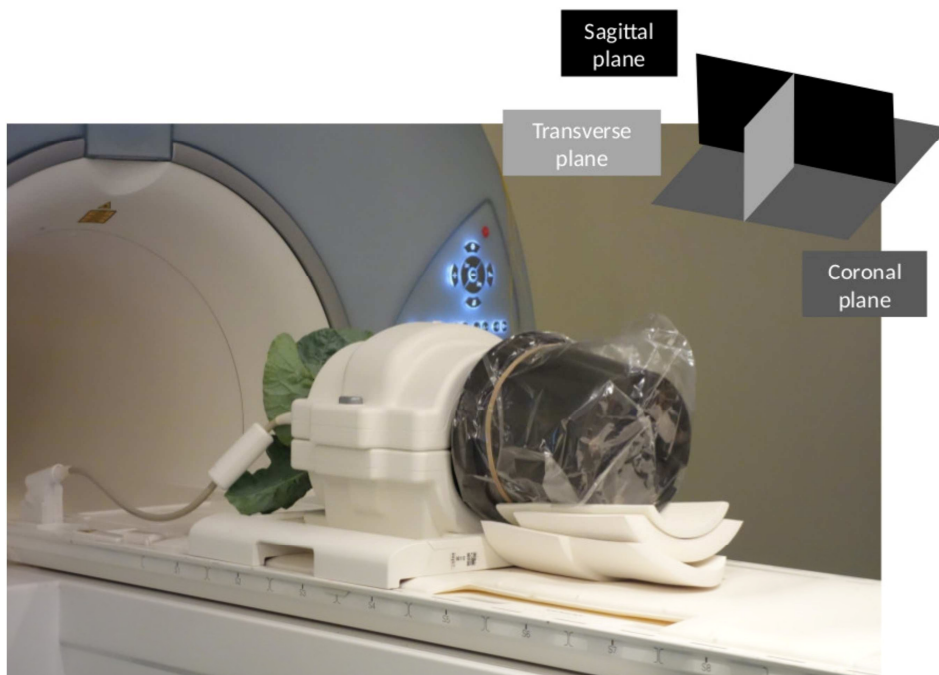


700

701

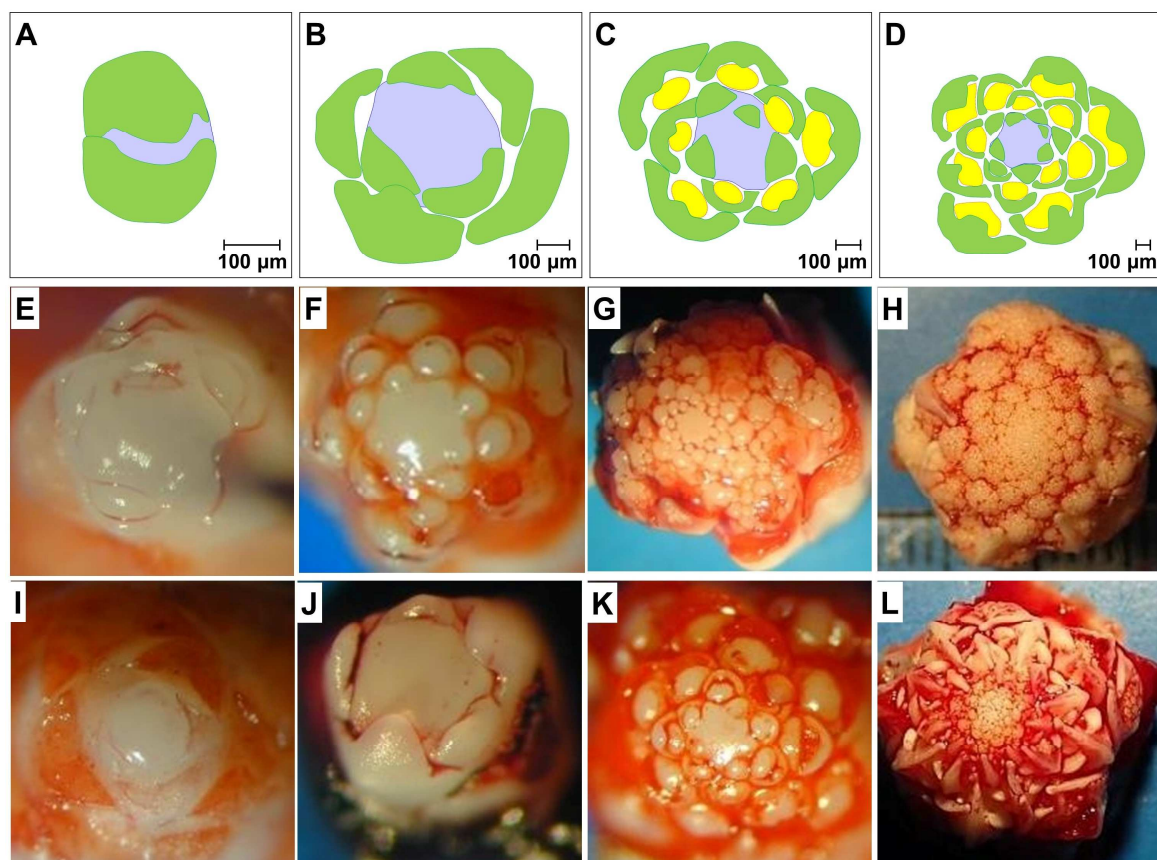
702

703 **Figure 3:** Example of one cauliflower plant placed in the "knee receiver coil" just before an
704 MRI acquisition.



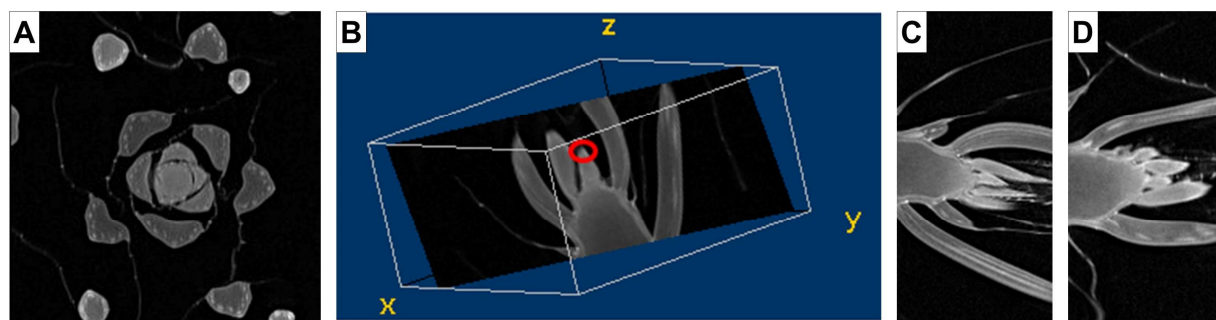
705
706
707

708 **Figure 4:** Floral induction at cauliflower apex illustrated by 4 developmental stages. Top line:
 709 schematic representation; middle line: examples in group H; bottom line: examples in group S.
 710 A,E,I: vegetative stage with only leaf scales (green/dark grey, plain line); B,F,J: curd-
 711 induction stage with enlargement of meristem center (blue/light grey, dashed line); C,G,K:
 712 curd-forming stage with round floral primordia (yellow/light grey, dashed line) initiated at the
 713 axil of each bract scale; D,H,L: curd-thickening stage with center only consisting of round
 714 floral primordia.



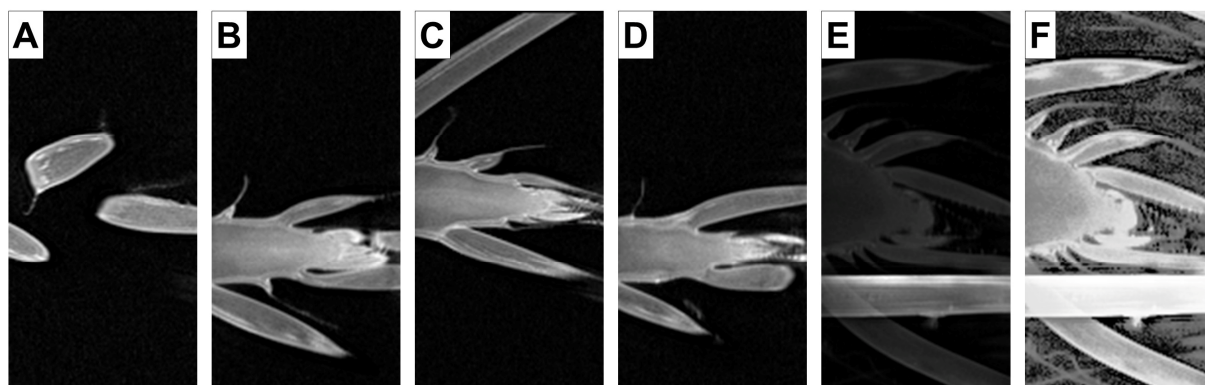
715
 716
 717
 718
 719

720 **Figure 5:** Examples of raw images in 2D planes and corresponding 3D reconstruction. A. raw
721 image from a sample on plane XY; B. 3D reconstruction with this sample's raw images by Fiji
722 [Schindelin et al. 2012]; C,D. middle slice on plane XZ and YZ. See Fig. 3 for a description of
723 the planes.



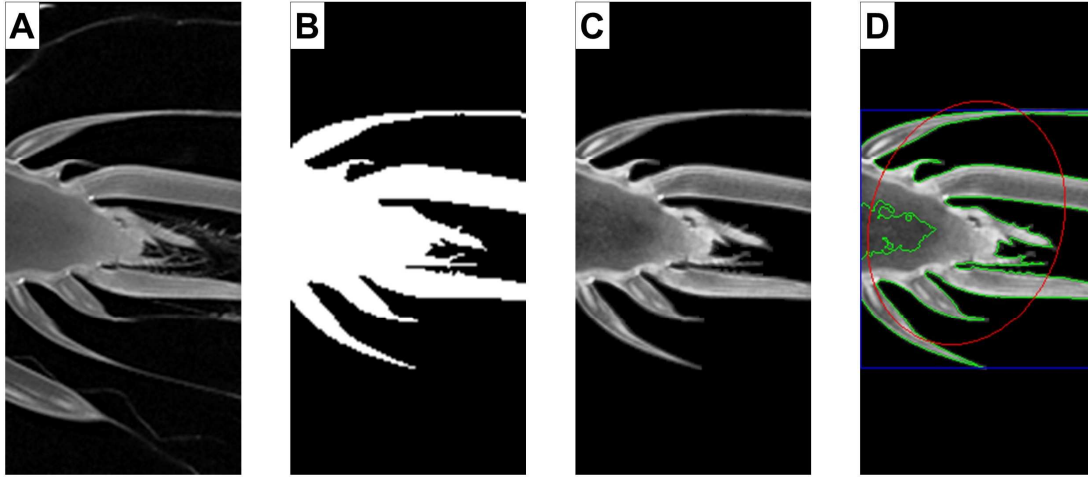
724
725
726
727
728

729 **Figure 6:** Illustrations for manual slice selection and histogram equalization A,B. missed
730 meristem in middle slices of a sample on plane XZ and YZ; C,D. best illustrative slices of the
731 same sample on plane XZ and YZ; E. best slice before histogram equalization; F. after
732 Histogram Equalization.



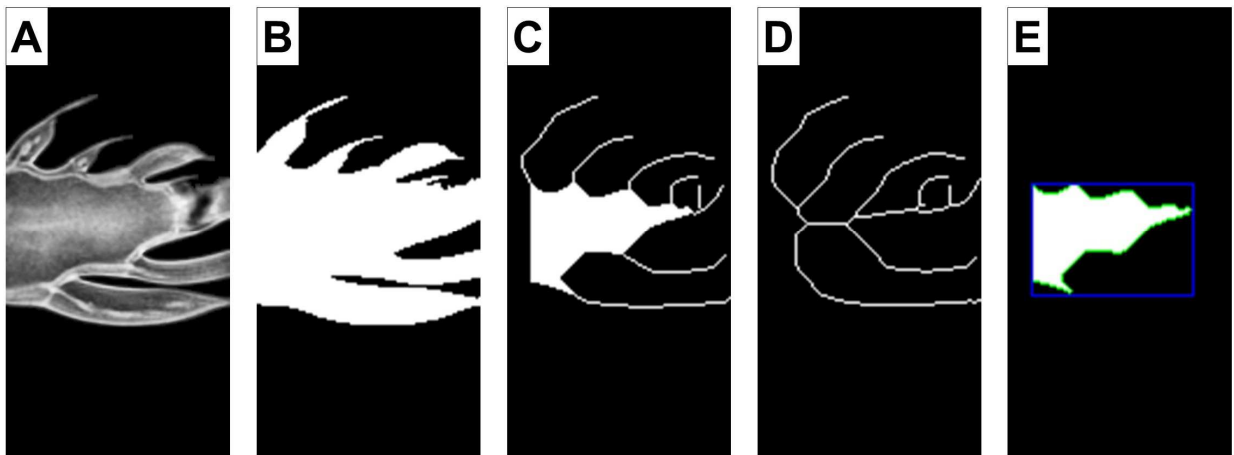
733
734

735 **Figure 7:** Extraction of the region of interest and resulting contour. A,B,C. illustration on
736 extraction of region of interest. D. image of contour where contour is marked in green,
737 rectangle in blue and ellipse in red.



738
739
740
741

742 **Figure 8:** Extraction of the image of skeleton from a ROI image, from A. to E.



743
744
745

746 **Appendix**

747

748 **Table 1:** The Jaccard Index (JI) of classifiers based on features from image of contour or/and
 749 skeleton by Linear Discriminant Analysis (LDA), Quadratic Discriminant Analysis (QDA),
 750 Partial Least Squares Discriminant Analysis (PLSDA). For comparison, on stage 234, the 2D
 751 CNN JI was 64,26 % and the 3D CNN JI was 45.52 %.

752

Data	Stage	Features	Size	H	S	LDA JI	QDA JI	PLSDA JI
2018 + 2019	1	image of contour	88	22	66	77.27%	79.55%	76.14 %
2018 + 2019	1	image of skeleton	88	22	66	82.95%	67.05%	82.95 %
2018 + 2019	1	image of contour + skeleton	88	22	66	79.55%	71.59%	78.41 %
2018 + 2019	2	image of contour	54	34	20	90.74%	83.33%	96.30 %
2018 + 2019	2	image of skeleton	54	34	20	55.56%	57.41%	55.55 %
2018 + 2019	2	image of contour + skeleton	54	34	20	77.78%	75.93%	81.48 %
2018 + 2019	3	image of contour	26	20	6	88.46%	73.08%	80.77 %
2018 + 2019	3	image of skeleton	26	20	6	69.23%	80.77%	76.92 %
2018 + 2019	3	image of contour + skeleton	26	20	6	80.77%	76.92%	76.92 %
2018 + 2019	4	image of contour	152	96	56	86.84%	82.89%	81.58 %
2018 + 2019	4	image of skeleton	152	96	56	86.18%	87.50%	85.53 %
2018 + 2019	4	image of contour + skeleton	152	96	56	88.82%	88.16%	89.47 %
2018 + 2019	234	image of contour	232	150	82	83.19%	82.76%	85.34 %
2018 + 2019	234	image of skeleton	232	150	82	77.59%	79.74%	76.293 %
2018 + 2019	234	image of contour + skeleton	232	150	82	84.05%	85.34%	84.91 %
2018	4'	image of contour	136	80	56	86.03%	85.29%	83.82 %
2018	4'	image of skeleton	136	80	56	83.82%	86.03%	86.03 %
2018	4'	image of contour + skeleton	136	80	56	88.24%	85.29%	92.65 %
2018	234'	image of contour	172	94	78	82.56%	78.49%	80.23 %
2018	234'	image of skeleton	172	94	78	75.00%	73.26%	71.51 %
2018	234'	image of contour + skeleton	172	94	78	79.65%	81.98%	83.72%

753

754

755