

Article

Modeling Chickpea Productivity with Artificial Image Objects and Convolutional Neural Network

Mikhail Bankin [†] , Yaroslav Tyrykin [†], Maria Duk, Maria Samsonova and Konstantin Kozlov ^{*} 

Mathematical Biology and Bioinformatics Lab, PhysMech Institute, Peter the Great St. Petersburg Polytechnic University, 195251 St. Petersburg, Russia; mikle.p.bankin@gmail.com (M.B.); tyrykinyarik@mail.ru (Y.T.); duk@mail.ioffe.ru (M.D.); m.g.samsonova@gmail.com (M.S.)

^{*} Correspondence: kozlov_kn@spbstu.ru; Tel.: +7-812-290-9642

[†] These authors contributed equally to this work.

Abstract: The chickpea plays a significant role in global agriculture and occupies an increasing share in the human diet. The main aim of the research was to develop a model for the prediction of two chickpea productivity traits in the available dataset. Genomic data for accessions were encoded in Artificial Image Objects, and a model for the thousand-seed weight (TSW) and number of seeds per plant (SNpP) prediction was constructed using a Convolutional Neural Network, dictionary learning and sparse coding for feature extraction, and extreme gradient boosting for regression. The model was capable of predicting both traits with an acceptable accuracy of 84–85%. The most important factors for model solution were identified using the dense regression attention maps method. The SNPs important for the SNpP and TSW traits were found in 34 and 49 genes, respectively. Genomic prediction with a constructed model can help breeding programs harness genotypic and phenotypic diversity to more effectively produce varieties with a desired phenotype.

Keywords: artificial image objects; climatic factors; genomic prediction; chickpea; GWAS; convolutional neural network



Citation: Bankin, M.; Tyrykin, Y.; Duk, M.; Samsonova, M.; Kozlov, K. Modeling Chickpea Productivity with Artificial Image Objects and Convolutional Neural Network. *Plants* **2024**, *13*, 2444. <https://doi.org/10.3390/plants13172444>

Academic Editor: Changsoo Kim

Received: 17 July 2024

Revised: 23 August 2024

Accepted: 28 August 2024

Published: 1 September 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The relevance of research in the field of agricultural crops is determined by the need to increase their productivity and resistance to adverse environmental conditions. One of the important crops is the chickpea, which plays a significant role in global agriculture due to its nutritional value and ability to fix nitrogen, thereby increasing soil fertility.

Chickpeas, being an accessible and rich source of nutrients, occupy an increasing share in the human diet every year. The chickpea is the second most widely grown food legume that is cultivated in more than 50 countries around the world, especially in West Asia and the Indian subcontinent. The chickpea provides nutritional nitrogen and high-quality protein for ≈15% of the world's population. Thus, over the past 12 years, the share of chickpea consumers in the United States has increased by 236%, reaching 4.5% of the entire population [1].

The main targets for breeding in chickpeas include *Fusarium* wilt [2]; increased resilience to drought, heat, and cold; increased seed nutrient density; reduced dependence on inputs; and resistance to biotic stress [3]. The application of omics technologies in breeding has proven effective in chickpeas. Joint analysis of the available phenotypic and genomic data identified candidate markers for many agronomic traits. Among the recent results, 27 marker–trait associations (MTAs) linked to yield-related traits and heat stress were identified in [4] using the BLINK model, and 11 SNPs associated with *Fusarium* wilt resistance, dispersed across the genome, were found in [5]; several marker–trait associations (MTAs) associated with the drought-linked traits were identified in [6]. Breeding strategies based on genomic prediction to enhance crop productivity have been proposed.

Genomic prediction (GP) aims to predict the phenotype of an organism given the data on single-nucleotide polymorphisms (SNPs) [7]. The plethora of methods for genomic prediction

can be classified in two groups. Linear methods such as BLUP perform well for additive traits. They model the phenotype as a function of contributions from different factors, e.g., individual markers, weather parameters, field conditions, etc. Dimensionality reduction methods are often used as a pre-processing step for genomic prediction methods [8]. On the other hand, nonparametric machine learning methods, e.g., Support Vector Machines, Random Forest, and Gradient Boosting Machine, can model nonlinear traits, providing tremendous flexibility to adapt to complicated associations between data and output [9].

The accuracy of prediction is affected by the quality and pre-processing of the phenotypic data, as well as the platform used to obtain genomic information, the population mating design, the intrinsic genetic architecture of the trait, the genetic structure of the population, how the genotype-by-environment interaction is dealt with, and the prediction method [10].

Among machine learning methods, convolutional neural networks (CNNs) provide the best ability to identify latent patterns or features from data and are best suited for image analysis [11,12]. Artificial Image Objects (AIOs) constitute a new concept for the representation of genomic data that can be used to encode large amounts of genomics data by considering individual genetic variants as pixels [13]. The advantages of AIOs are convenient straightforward visualization, compactness, and ability to apply vast number of techniques developed for image analysis and classification [14], particularly for the CNN [15]. The need for diverse methods for data representation and processing is particularly important due to the great increase in the numbers of genomic data such as *Cicer* super-pangenome [16]. Consequently, AIOs could be used by CNNs on regression and classification tasks [17].

Application of the CNN to the processing of AIOs makes it possible to calculate and visualize the impacts of different factors to the final model solution. Recently, increasing attention has been drawn to the internal mechanisms of convolutional neural networks, as well as the reason why the network makes specific decisions [18]. Several techniques have been introduced that include perturbation- and backpropagation-based approaches [19], gradient-based algorithms [20], and class activation maps [18]. A saliency map represents the spatial support of a particular class in a given image [21].

Though the application of existing methods may seem to be straightforward, the development of new approaches will broaden the range of available opportunities for the prediction of important plant characteristics. Modern machine learning algorithms overcome some of the shortcomings of classical methods, especially in the ability to model complex dependencies between data and output, but present new challenges, such as the selection of adequate data representation and model architecture that are addressed in this work.

The aim of the present research was to develop a model for chickpea productivity traits in the available dataset and to extract the most important features that influence the model solution.

The main contributions of this work are the following:

- A methodology was proposed that combines AIOs and modern ML algorithms;
- A model based on AIOs and a CNN was developed for the prediction of chickpea productivity traits using SNPs;
- The impacts of SNPs on the model solution were evaluated.

The developed approach can help breeding programs harness genotypic and phenotypic diversity to more effectively produce varieties with a desired phenotype.

2. Related Work

Machine learning methods have been used to increase the statistical power of GWAS [22], to detect epistatic interactions, to improve the polygenic risk estimate obtained using GWAS procedures [23] and for post-processing the results of GWAS analysis [24]. Recently, improved Random Forest (RF) [25] methods have been proposed and applied to GWASs, such as the T-Tree method and the ts-RF method [26], which optimize tree node separation rules;

Szymczak (2016) [27] redefined the method for calculating importance scores; in [28], a hybrid RF algorithm was proposed.

The work of [29] was said to be the first study to apply a saliency map for a GWAS, and the missing values were treated as a new genotype resulting in four binary values for each SNP in soybean. The saliency value of each genotype was calculated as the maximum absolute value of gradients among those four coding channels; the population median value was used as a measurement of the SNP contribution.

Various methods for the interpretation of CNN models have been proposed recently. Class activation maps provide visual explanation for a single input [20,30] but are architecture-sensitive. Gradient-weighted Class Activation Mapping (Grad-CAM) uses the gradients of any target concept flowing into the final convolutional layer to produce a coarse localization map highlighting the important regions in the image for predicting the concept [20]. Score-CAM, unlike previous class-activation-mapping-based approaches, gets rid of the dependence on gradients by obtaining the weight of each activation map through its forward passing score on the target class; the final result is obtained by a linear combination of weights and activation maps [18]. Grad-CAM++ [30], the modification of Grad-CAM [20], generalizes CAM to models without global pooling layers. LayerCAM [31] may generate reliable class activation maps from combinations of class activation maps from different layers of the CNN.

DeepFeature, proposed in [32,33], was designed to transform omics data into a form that is optimal for fitting a CNN model. The input vector is transformed into a matrix using t-SNE, kernel PCA, PHATE, or UMAP, and the smallest rectangle containing all the elements is found using the convex hull algorithm; rotation is performed to align the image, and Cartesian coordinates are converted to pixel indices.

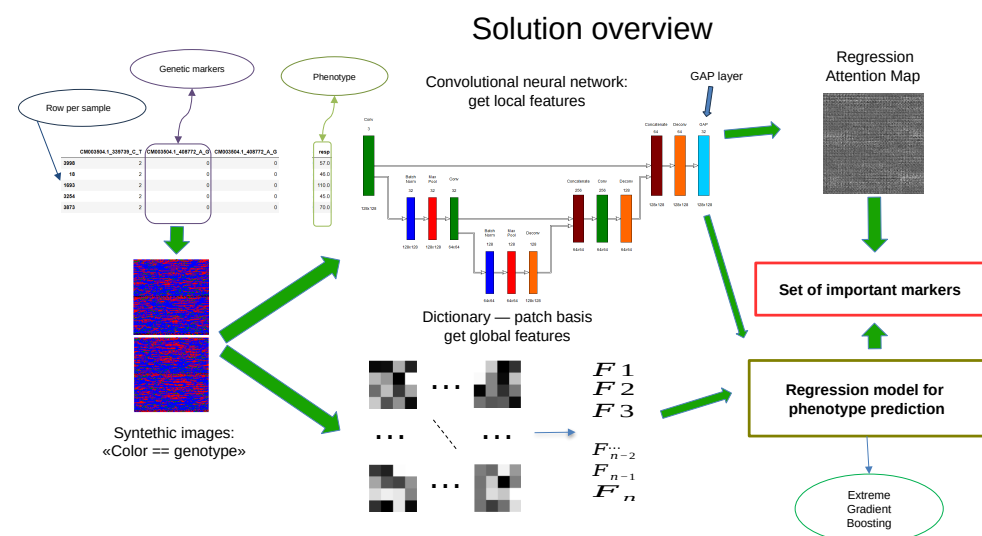
3. Materials and Methods

3.1. The Overview

The methodology proposed in this work consists of several steps (see Figure 1):

1. Construction of artificial images for each accession by encoding information on the SNP values and climatic factors for limited period of time;
2. Building convolutional neural network for local feature extraction;
3. Dictionary learning and sparse coding for extraction of global features;
4. Construction of an extreme gradient boosting model for prediction of chickpea traits,
5. Evaluation of importance of input data for model prediction using the regression activation mapping technique.

These steps are described below.



3.2. Plant Material

A total of 407 chickpea *Cicer arietinum* samples collected in Ethiopia, Lebanon, Morocco, Turkey, India, Uzbekistan, and the Mediterranean region were phenotyped at the VIR Kuban experimental station in 2016. During the vegetative period, 36 phenological, morphological, agronomical, and biological descriptors were measured. Details on the phenotyping experiments, genotyping, and subsequent analysis resulted in 6642 SNPs that were presented in recent manuscripts [34].

In this work, two productivity traits were modeled, namely, the thousand-seed weight (TSW), which ranges from 110 to 440 g (see Figure 2), and the number of seeds per plant (SNpP), which ranges from 0 to 88 seeds (see Figure 3).

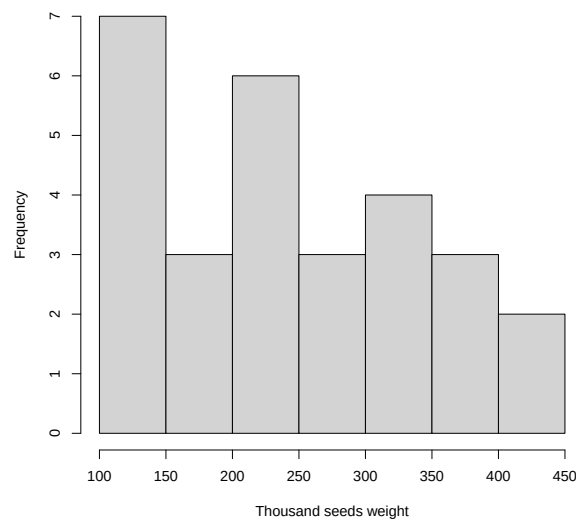


Figure 2. Histogram for TSW trait.

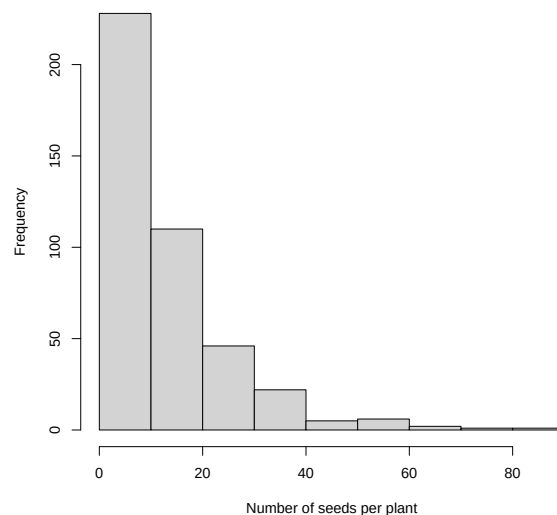


Figure 3. Histogram for SNpP trait.

3.3. Artificial Image Objects

Artificial image objects were used to encode information on V_g genomic features for each accession.

As the dimensions of AIOs may be selected arbitrarily, it was decided to set the number of rows and columns equal to 128 pixels. AIO $I(x, y)$ can be represented as a matrix that is sequentially filled with the values of features from left to right in row-first manner; the list of features was reused from the start in order to fill the whole matrix. Each pixel

value combines three channels, R , G , and B , for three pseudo-colors—red, green, and blue, respectively.

The value f_g of SNP with index k was converted to a pixel value $i_g(x, y)$ according to (1):

$$R = \begin{cases} 1, & f_g = 0 \\ 0, & f_g \neq 0 \end{cases} \quad G = \begin{cases} 1, & f_g = 1 \\ 0, & f_g \neq 1 \end{cases} \quad B = \begin{cases} 1, & f_g = 2 \\ 0, & f_g \neq 2 \end{cases} \quad (1)$$

The artificial image objects were constructed for all accessions. AIOs provide a convenient visualization for the data (see Figure 4). The colors of pixels in the AIOs are defined by Equations (1) for genomic factors.

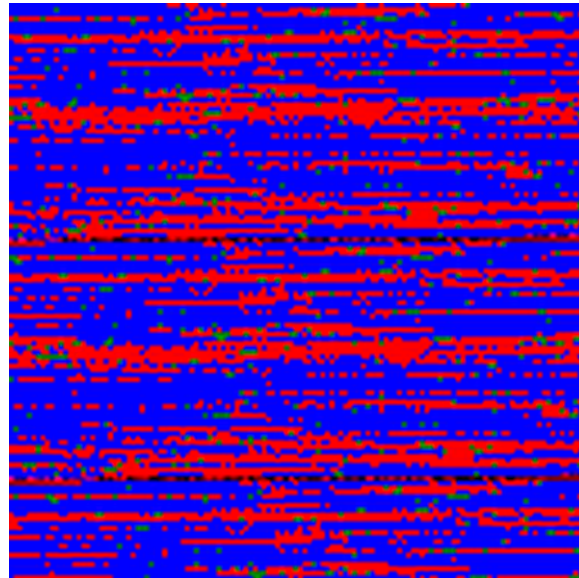


Figure 4. Example AIO. The size of the image is 128×128 px; here, the image is enlarged. Each colored square corresponds to one pixel. The color of each pixel is obtained by (1) for genomic factors.

3.4. Dictionary Learning and Sparse Coding

To account for the inter-relations between accessions in the dataset due to the genetic similarity that can impact the model performance, the additional features were extracted from the set of AIOs using the online dictionary learning and sparse coding approach [35].

The individual images are combined into one, which allows the algorithm to analyze all available features from all images and contributes to their closer integration. The combined image is then split into square blocks of the selected small dimension, for example, 16 by 16 pixels. Each image block can be represented as a weighted sum of the selected number, e.g., 16, of template images (called atoms) stored in a dictionary. In the training phase, template images are obtained from a set of blocks of real images, with the k -nearest neighbors method used to find similar templates and the singular value decomposition (SVD) method used to speed up the convergence of the iterative dictionary construction process.

The non-zero coefficients for each dictionary atom are distributed according the log-normal law, so the fitted means and standard deviations are used to calculate the vector of the microfeatures. The frequencies of atoms constitute the vector of the macrofeatures [36]. The concatenation of these two vectors was used as the numerical feature vector in the model.

3.5. Convolutional Neural Network

The model for phenotype prediction was built in the form of a convolutional neural network (see Figure 5) that has been improved in comparison to [17]. The CNN takes artificial image objects (img_input in Figure 5) and the numerical feature vector (dict_input in Figure 5) as the first and the second inputs, respectively. The size of the images was fixed

at 128 by 128 pixels. Since the AIO contains three color channels, each filter is a collection of three kernels. Each kernel slides along the corresponding image channel; the results are combined into one feature map. The size of the filter kernel of each convolutional layer (Conv2D type) was found by adapting to the available experimental data. The values of the weights of the kernels were the learning parameters of the model. Each convolutional layer was followed by a subsampling layer (MaxPooling2D type), the purpose of which was to reduce the dimension of maps in order to enlarge features. Such filtering helps, among other things, to avoid overfitting. The formation of a new feature map is based on the Max Pooling operation, which is performed by selecting the maximum value from a subsample of a given size. The result is a set of 128 channels of 32-by-32-pixel feature maps. Next, using the operations of deconvolution, combining with intermediate feature maps to prevent loss of information, and using convolution, 32 channels of feature maps of the original size 128 by 128 pixels were obtained.

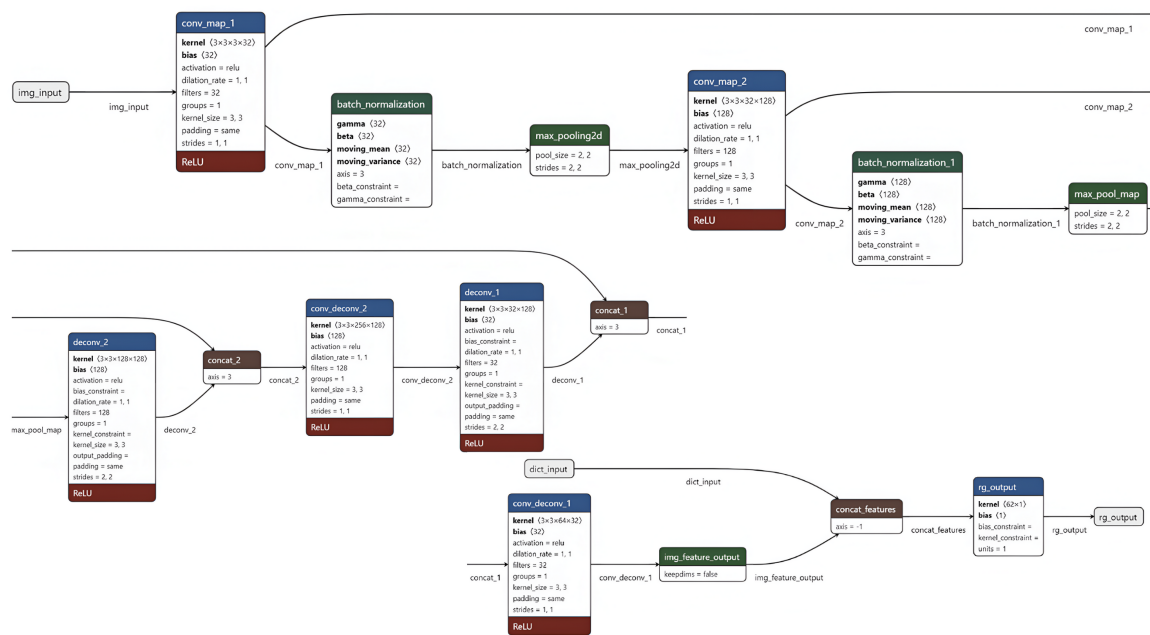


Figure 5. The architecture of CNN.

At the last stage, all feature maps were combined using a global average pooling layer (GAP) that transformed the maps into a feature vector, which was combined with the numerical feature vector. The combined vector was used to obtain a solution using a regression extreme gradient boosting model. The mean absolute error was used as a loss function. The neural network and regression model were trained simultaneously to obtain a consistent solution.

The network architecture includes 536,575 parameters and was implemented using the TensorFlow v2.12 and Keras v2.12 packages using a functional interface in the Python programming language.

3.6. Impacts of Different Factors to the Model Solution

Attention maps are a widely used method for explaining the results of convolutional neural networks. Classification and regression tasks use attention maps based on the values of network function gradients or the convolutional layers. In this work, we adapted the dense regression attention maps proposed in [37,38].

Attention maps were computed for input images obtained from genomic data of individual accessions and were of the same size. Thus, the pixel intensity characterizes the importance of a particular feature—SNP—for solving the model for a given sample. Consequently, in order to identify the most important features, it is necessary to determine

those for which the average pixel intensity significantly differs from the averages for other features. We used Dunnett's criterion [39], a well-known statistical test for many-to-one comparisons, to compare pixel averages.

4. Results

4.1. Dictionary Learning

The genetic similarities between accessions may affect the accuracy of the genomic prediction if not accounted in the model structure. The statistical parameters of the AIO decomposition coefficients according to the learned dictionary were used as the supplemental numerical features that characterize the population structure present in the dataset.

To optimize the solution, online dictionary learning method and the orthogonal matching pursuit method were adapted. The advantages of these algorithms are the reduced memory consumption during calculations and better accuracy. Several dictionaries were built for the dataset for error values of 0.005, 0.01, and 0.02 and for decomposition lengths of 10 and 20.

The dictionary with 16 atoms and an error value of 0.01 was selected, as it had the smallest number of missing coefficients.

4.2. Model for Number of Seeds per Plant

The model for the SNpP was trained using AIOs of 366 and 41 accessions for 10-fold cross-validation and control, respectively. The validation split parameter was set to 20%, and training was performed for 150 epochs using the mean absolute error as the loss function. The loss function curves for the training and validation subsets achieved approximately the same level, so the learning process was considered to be converged (see Figure 6). The model with the maximal accuracy for the validation set was considered to be the best and was selected for further investigation. The best model predicted the number of seeds per plant for the control dataset with a high accuracy of 84% (see Figure 7). This result is good for the task of predicting phenotypes using SNPs.

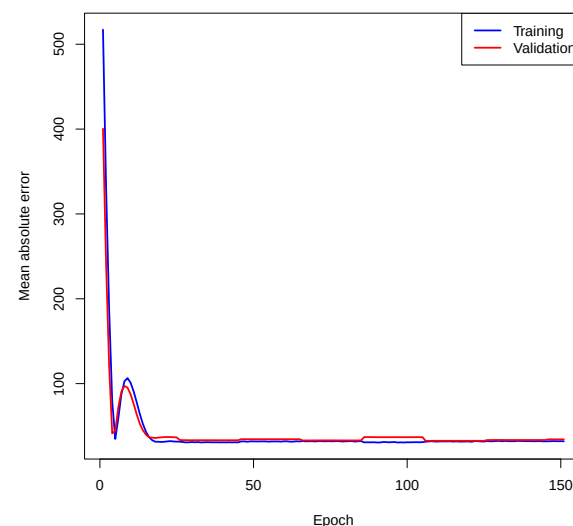


Figure 6. Convergence for SNpP trait.

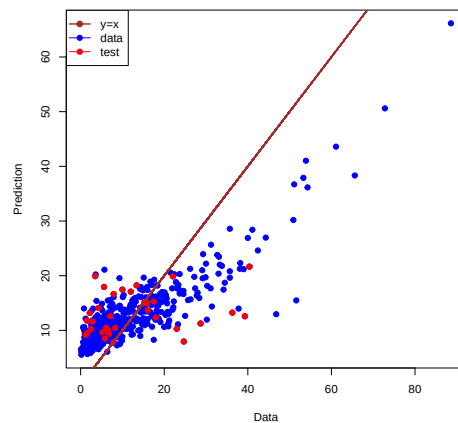


Figure 7. Comparison of measured and predicted number of seeds per plant. The data points used for training are marked with blue circles, and those from test set are drawn as red dots. The straight line represents the exact correspondence. The model accuracy was $a = 84\%$.

4.3. Important Features for Number of Seeds per Plant

The attention maps were computed for each individual accession in order to identify SNPs with the largest impact on trait prediction. The individual map (see Figure 8) contains pixels of a high range of intensities, with several comparatively lighter areas containing potentially important SNPs.

A total of 99 SNPs were selected for the SNpP trait using the approach described in Section 3.6.

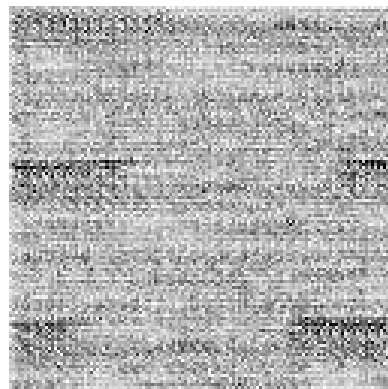


Figure 8. Example attention map for SNpP trait for individual accession. The intensity differences were increased for visualization purposes.

4.4. Functional Analysis of Identified SNPs for SNpP

For the SNpP, 34 out of 99 SNPs were located within known genes or within 1 kb flanking regions, and they most likely tag candidate casual genes (Supplementary Table S2). The functions of some of them are known.

The statistical significance of the difference in phenotype means was checked with the Student's *t*-test between samples that were reference homozygous, alternative homozygous, and heterozygous for the identified SNPs. The difference in the SNpP was significant (p -value < 0.05) for 23 out of 34 SNPs within genes and 1 kb flanking regions and for 28 more SNPs out of the remaining 65 identified for this phenotype (Supplementary Table S3).

The statistical significance of the difference in phenotype means was checked with a Student's *t*-test between samples that were reference homozygous, alternative homozygous, and heterozygous for the identified SNPs. The difference for the TSW phenotype was significant (p -value < 0.05) for 8 out of 49 SNPs within genes and 1 kb flanking regions and for 3 more SNPs among the rest of the identified 50 SNPs. The difference in the SNpP was

significant (p -value < 0.05) for 23 out of 34 SNPs within genes and 1 kb flanking regions and for 28 more SNPs out of the rest of the 65 identified for this phenotype (Supplementary Table S3).

For the SNpP trait, an SNP (Chr1:7478974) was found within the *Ca_08000* gene encoding the biotin carboxyl carrier protein of acetyl-CoA carboxylase, which plays crucial roles in fatty acid metabolism [40].

An SNP (Chr1:7497743) was located 305 bp upstream of the *Ca_08003* genes for subunit 10 of the ER membrane protein complex, which inserts newly synthesized proteins into membranes [41] and performs essential functions, including protein and lipid synthesis [42]. An SNP (Chr1:26966675) was found downstream of the *Ca_20295* gene for the SEH1 protein. In *Arabidopsis thaliana*, SEH1 is implicated in immunity-related mRNA export [43]. An SNP (Chr1:19804162) was found within the *Ca_08855* gene encoding serine/threonine-protein kinase PCRK1-like required for plant immunity [44].

Three SNPs (Chr1:6354038; Chr1:6354059; and Chr1:6354082) were found within the *Ca_07897* gene for DEAD-box ATP-dependent RNA helicase 8. In *Arabidopsis thaliana*, this helicase modulates the formation of D-bodies, which are membraneless organelles that are made from liquid droplets of proteins and nucleic acids [45].

An SNP (Chr1:26891914) was located within the *Ca_20291* gene encoding zinc-finger homeodomain (ZF-HD) protein 2-like. All members of the subfamily of the ZF-HD homeobox genes are expressed predominantly or exclusively in floral tissue, indicating a likely regulatory role during floral development [46]. An SNP (Chr1:7592329) was found 213 bp downstream of the *Ca_08013* gene for AP2-like ethylene-responsive transcription factor AIL1 that is involved in the specification of meristematic or division-competent states [47].

An SNP (Chr1:31369599) was found within the *Ca_21855* gene controlling the long-distance transport of glucosinolates, which are defense compounds. The homologs of this gene in *Arabidopsis* control the loading of glucosinolates from the apoplasm into the phloem and into the seeds [48]. An SNP (Chr1:8130166) was found within the *Ca_08059* gene, which is the Glyma18g48580 homologue that encodes the subtilisin-like protease involved in plant immunity [49]. The *Ca_25478* gene, which contains an (Chr1:34314586) SNP, is a homologue of the *Arabidopsis thaliana* LEAF RUST 10 DISEASE-RESISTANCE (AtLRK10L1) gene. Mutation in this gene results in an abscisic acid (ABA)-insensitive phenotype in seed germination and seedling growth, as well as in reduced tolerance to drought stress [50].

An SNP (Chr1:27024341) was located within the *Ca_20299* gene for FAR1-RELATED SEQUENCE 5-like protein, which is a homologue in *Arabidopsis thaliana* that is essential for phytochrome A controlled far-red responses [51]. In addition to its essential role in light signaling, FAR1 also plays diverse regulatory roles in plant growth and development, including clock entrainment, seed dormancy and germination, senescence, chloroplast formation, branching, and flowering and meristem development [52].

An SNP (Chr1:6162693) was found within the *Ca_00701* gene for 1-aminocyclopropane-1-carboxylate oxidase, which catalyses the final step in ethylene biosynthesis—an important regulator of many developmental and physiological processes such as seed dormancy, germination, vegetative growth, flowering, climacteric fruit ripening, and senescence [53].

4.5. Model for Thousand-Seed Weight

The model training for the TSW trait prediction was performed using 366 accessions for 10-fold cross-validation. The loss function and training parameters were the same as for the SNpP model, i.e., 150 epochs and a 20% validation subset. The convergence process is shown in (Figure 9). The best model was selected that had the maximal accuracy for the validation set. The best model predicted the TSW trait for the control dataset with a high accuracy of 85% (see Figure 10).

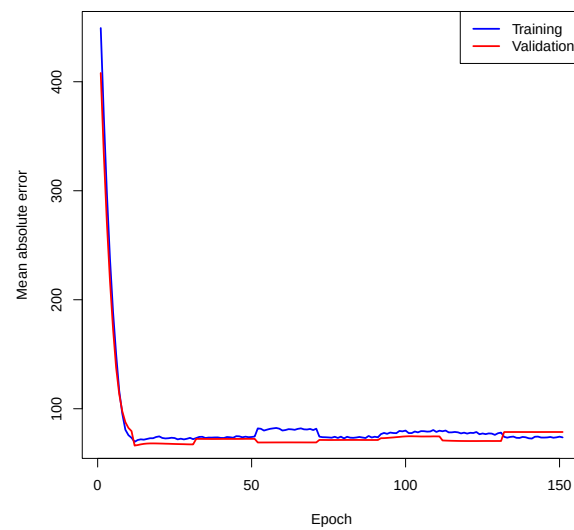


Figure 9. Convergence for TSW trait.

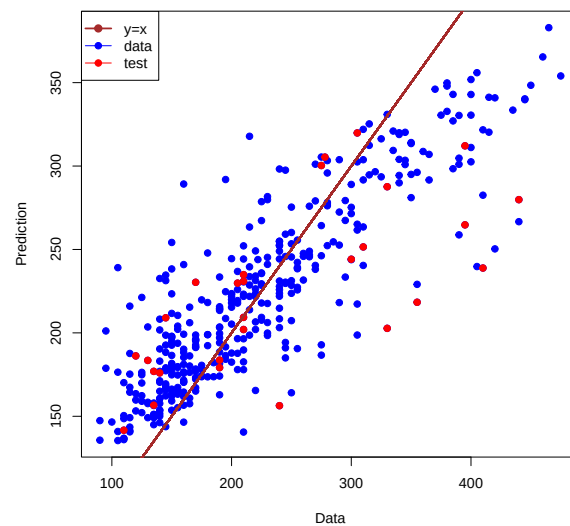


Figure 10. Comparison of measured and predicted thousand-seed weight. The data points used for training are marked with blue circles, and those from test set are drawn as red dots. The straight line represents the exact correspondence. The model accuracy was $a = 85\%$.

4.6. Important Features for Thousand-Seed Weight

In order to find the most important SNPs, the attention maps (see Figure 11) were computed for each individual accession.

The individual attention map for this model (Figure 11) looked similar to that for the previous one (Figure 8). The gray levels of the pixels varied greatly, and several comparatively brighter islands can be identified that correspond to potentially important SNPs.

Following the same procedure described in Section 3.6, 99 SNPs were found for the TSW trait.

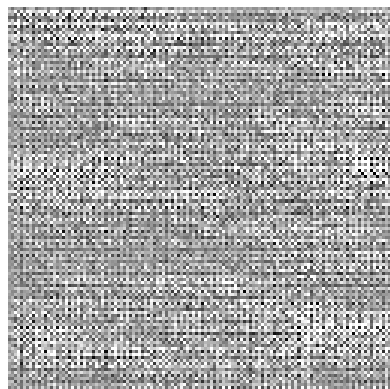


Figure 11. Example attention map for TSW trait for individual accession, the intensity was enhanced for visualization similar to Figure 8.

4.7. Functional Analysis of Identified SNPs for TSW

For the TSW trait, 49 out of 99 SNPs were located within known genes or within 1 kb flanking regions and most likely tag candidate casual genes (Supplementary Table S1).

The statistical significance of the difference in phenotype means was checked with a Student's *t*-test between samples that were reference homozygous, alternative homozygous, and heterozygous for the identified SNPs. The difference for the TSW phenotype was significant (p -value < 0.05) for 8 out of 49 SNPs within genes and 1 kb flanking regions and for 3 more SNPs among the rest of the identified 50 SNPs (Supplementary Table S3).

For the TSW trait, the identified SNPs were found in known genes involved in the biosynthesis and transport of nutrients and adaptation to unfavorable environmental conditions. Two SNPs (Chr4:38305135; Chr4:38305142) were within the *Ca_13125* gene encoding the diacylglycerol O-acyltransferase 1C-like protein DGAT1. DGAT is considered to be a key enzyme in the conversion of diacylglycerol (DAG) to TAG. Triacylglycerols (TAGs) are the major seed storage lipids, providing carbon and energy reserves to support seedling growth during germination [54]. TAGs are involved in mediating pollen development and sexual reproduction in many plant species (DGAT1) [52]. A study conducted in rapeseed plants (*Brassica napus* L.) showed that inhibition of DGAT1 resulted in reduced seed oil content and germination rates, as well as severe developmental abnormalities [55].

The SNP Chr1:1457542 was found within *Ca_00178*, which encodes the oligopeptide transporter 4 known as OPT4. In *Arabidopsis thaliana* plants, OPT4 transporter proteins are responsible for loading peptides into the plant vasculature [56].

Three SNPs (Chr2:35757170; Chr2:35759932; and Chr2:35756260) were found in the *Ca_09749* gene, encoding the NF-kappa-B-activating protein Nkap. MSA2, an orthologue of human Nkap in *Arabidopsis thaliana*, is known to be involved in the regulation of rRNA synthesis [57].

The *Ca_08059* gene contained three SNPs (Chr1:8130179; Chr1:8131436; and Chr1:8131450), which encodes a protein from the subtilase family—subtilisin-like protease Glyma18g48580. Studies on soybean *Glycine max* L. showed the role of this protein in its interaction with the GmSubPep peptide and subsequent activation of defenses in response to the presence of pathogens [58].

Our study identified an SNP known as Chr6:10613546 located in the *Ca_08536* gene encoding the E3 ubiquitin-protein ligase SINAT2 protein. In *Hordeum vulgare* L. SINAT2 was identified as a candidate gene involved in salt response [59]. Salinity limits chickpea growth and productivity by interfering with various physiological and metabolic processes [60]. On average, salinity can reduce chickpea plant growth rates by 20%, plant heights by 15%, and shoot biomass levels by 28%. It also causes unfilled pods and reduces the seed number and yield by 16% and 32%, respectively [61].

The SNPs Chr1:2511223 and Chr1:2511238 are located in the *Ca_00314* gene, encoding the acetolactate synthase 3 AHAS3. It is known that a point mutation in the chickpea AHAS1 gene results in an amino acid change from Ala205 to Val205, which makes chickpeas

resistant to the herbicide imidazolinone [62]. This gene was found in the study [63] to be associated with the seeds per plant and days to maturity traits (see Suppl.tab.14 in [63]).

An SNP (Chr1:5757141) was found in the *Ca_00667* gene encoding the glutathione-dependent enzymes glyoxalase I (GLX1) gene—the activity of which plays a critical role in the detoxification of methylglyoxal [64].

5. Discussion

In this work, a methodology was proposed that combines AIOs to encode information on genomic features—6642 SNPs for each accession, a CNN, dictionary learning and sparse coding for feature extraction, extreme gradient boosting for phenotype prediction, and dense regression attention maps for the identification of the most important genomic factors. This approach took into account the properties of convolutional neural networks.

Two main phenotypic traits were taken as productivity indicators for modeling: the thousand-seed weight (TSW) and number of seeds per plant (SNpP). The models were trained using 10-fold cross-validation and reached an acceptable accuracy of 84% and 85% for the TSW and SNpP, respectively.

For each trait, 99 SNPs were selected as the most important for the model solution, of which 34 (for SNpP) and 49 (for TSW) were located within known genes or within 1 kb flanking regions and most likely tag candidate causal genes. The functions of some of them are known.

The difference in phenotype means was statistically significant with the Student's *t*-test (p -value < 0.05) employed between samples that were reference homozygous, alternative homozygous, and heterozygous for the identified SNPs for 8 out of 49 SNPs and for 23 out of 34 SNPs for the TSW and SNpP traits, respectively.

Two SNPs (Chr1:2511223; Chr1:2511238) found in this study to be important for TSW are located at the distance 85 bp and 70 bp, respectively, from the SNP (Chr1:2511308) that was found by FarmCPU to be associated with the seeds-per-plant trait in [63]. The SNP (Chr4:44945463) from this study is located 744 bp away from the SNP (Chr4:44946207) that was found by FarmCPU to be associated with plant height in [63].

The accuracy of genomic prediction achieved using the developed approach suggests the need to explore the phenotypic performance of the genotypes of interest in different environments. One of the important implications of the present research is that the application of advanced machine learning methods for the prediction of phenotypic traits from genomic data may be computationally expensive but makes it possible to find new potential candidate genes that can be further investigated for use in breeding programs.

The main limitation of the proposed approach is the need to encode the genomic data in AIOs and compute the sparse representation according to the dictionary. It should be noted that this step needs to be done once for the dataset, and the results are then used for all traits. As the proposed model predicts the phenotype from features extracted from the images, the encoding scheme that maps the genomic data to pixel intensities may influence the model performance. The three-color channel scheme utilized in this work may not be appropriate in all cases, for example, for SNPs with more allele combinations.

6. Conclusions

The proposed methodology efficiently encoded genomic factors as AIOs and constructed a model based on dictionary learning and sparse coding, CNN, and extreme gradient boosting that predicted the seeds number per plant and thousand-seed weight values of accessions from a chickpea dataset with an acceptable accuracy of $\approx 85\%$. In the context of chickpea breeding, our approach can help to more effectively produce varieties with desired traits; however, further studies are necessary to provide new valuable and unbiased information for assessing its methodological potential.

The most important factors that influence the model solution were identified using dense regression attention maps. The important SNPs were found in known genes or within 1 kb flanking regions and most likely tag candidate causal genes the functions of some

of which that are known. The difference in phenotype means was statistically significant with the Student's *t*-test values yielded for a fraction of these SNPs, which supports the importance of these factors for traits of interest.

Future research in the application of artificial image objects and machine learning methods regarding the identification of important SNP markers and the construction of predictive models of important agronomic traits will include the optimization of the layout of factors in AIOs. The SNPs that are not close to each other in terms of genomes may be placed to the same part of the AIO and picked up together by the convolution kernel. Our future research will also include model verification with independent datasets.

Supplementary Materials: The following are available online at <https://www.mdpi.com/article/10.3390/plants13172444/s1>: Table S1: SNP identified in for TSW located in gene body (GB) or 1 kb flanking regions; Table S2: SNP identified in for SNpP located in gene body (GB) or 1 kb flanking regions; Table S3: results for statistical significance of difference in phenotype means for different allele combinations.

Author Contributions: Conceptualization, methodology, M.S. and K.K.; software, Y.T. and K.K.; validation, M.D., M.S. and K.K.; formal analysis, investigation, Y.T., M.B. and K.K.; resources, data curation, M.S. and M.B.; writing—original draft preparation, review and editing, Y.T., M.B., M.D., M.S. and K.K.; visualization, Y.T., M.D. and K.K.; supervision, M.S.; project administration, M.S.; funding acquisition, M.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Russian Science Fund grant number 22-46-02004 (Method development and the analysis of TSW trait, sect.1-3, 4.1, 4.5-4.7, 5, 6) and by the Ministry of Science and Higher Education of the Russian Federation as part of the World-class Research Center program: Advanced Digital Technologies contract No. 075-15-2022-311 dated 20 April 2022 (the analysis of SNpP trait, sect.4.2-4.4).

Data Availability Statement: The data for this publication is available at Zenodo <https://doi.org/10.5281/zenodo.12755678>.

Acknowledgments: We thank Vitaly Gursky, Svetlana Surkova, Anton Logachev, Nelly Gallyamova and Denis Gubrienko.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Rehm, C.D.; Goltz, S.R.; Katcher, J.A.; Guarneiri, L.L.; Dicklin, M.R.; Maki, K.C. Trends and Patterns of Chickpea Consumption among United States Adults: Analyses of National Health and Nutrition Examination Survey Data. *J. Nutr.* **2023**, *153*, 1567–1576. [CrossRef]
2. Laxuman, C.; Naik, Y.D.; Desai, B.K.; Kenganal, M.; Patil, B.; Reddy, B.S.; Patil, D.H.; Chakurte, S.; Kuchanur, P.H.; K, S.K.; et al. Development and evaluation of *Fusarium* wilt-resistant and high-yielding chickpea advanced breeding line, KCD 11. *Plant Genome* **2024**, *17*, e20460. [CrossRef] [PubMed]
3. von Wettberg, E.J.; Chang, P.L.; Başdemir, F.; Carrasquilla-Garcia, N.; Korbu, L.B.; Moenga, S.M.; Bedada, G.; Greenlon, A.; Moriuchi, K.S.; Singh, V.; et al. Ecology and genomics of an important crop wild relative as a prelude to agricultural innovation. *Nat. Commun.* **2018**, *9*, 649. [CrossRef] [PubMed]
4. Danakumara, T.; Kumar, N.; Patil, B.S.; Kumar, T.; Bharadwaj, C.; Jain, P.K.; Nimmy, M.S.; Joshi, N.; Parida, S.K.; Bindra, S.; et al. Unraveling the genetics of heat tolerance in chickpea landraces (*Cicer arietinum* L.) using genome-wide association studies. *Front. Plant Sci.* **2024**, *15*, 1376381. [CrossRef]
5. Alsamman, A.M.; H. Mousa, K.; Istanbuli, T.; Abd El-Maksoud, M.M.; Tawkaz, S.; Hamwieh, A. Unveiling the genetic basis of *Fusarium* wilt resistance in chickpea using GWAS analysis and characterization of candidate genes. *Front. Genet.* **2024**, *14*, 1292009. [CrossRef] [PubMed]
6. Istanbuli, T.; Nassar, A.E.; Abd El-Maksoud, M.M.; Tawkaz, S.; Alsamman, A.M.; Hamwieh, A. Genome-wide association study reveals SNP markers controlling drought tolerance and related agronomic traits in chickpea across multiple environments. *Front. Plant Sci.* **2024**, *15*, 1260690. [CrossRef]
7. Meuwissen, T.H.E.; Hayes, B.J.; Goddard, M.E. Prediction of Total Genetic Value Using Genome-Wide Dense Marker Maps. *Genetics* **2001**, *157*, 1819–1829. [CrossRef]
8. Manthena, V.; Jarquín, D.; Varshney, R.K.; Roorkiwal, M.; Dixit, G.P.; Bharadwaj, C.; Howard, R. Evaluating dimensionality reduction for genomic prediction. *Front. Genet.* **2022**, *13*, 958780. [CrossRef]

9. Montesinos-López, O.A.; Montesinos-López, A.; Pérez-Rodríguez, P.; Barrón-López, J.A.; Martini, J.W.R.; Fajardo-Flores, S.B.; Gaytan-Lugo, L.S.; Santana-Mancilla, P.C.; Crossa, J. A review of deep learning applications for genomic selection. *BMC Genom.* **2021**, *22*, 19. [\[CrossRef\]](#)
10. de los Campos, G.; Hickey, J.M.; Pong-Wong, R.; Daetwyler, H.D.; Calus, M.P.L. Whole-Genome Regression and Prediction Methods Applied to Plant and Animal Breeding. *Genetics* **2013**, *193*, 327–345. [\[CrossRef\]](#)
11. Pook, T.; Freudenthal, J.; Korte, A.; Simianer, H. Using Local Convolutional Neural Networks for Genomic Prediction. *Front. Genet.* **2020**, *11*, 561497. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Montesinos-López, O.A.; Montesinos-López, A.; Mosqueda-Gonzalez, B.A.; Montesinos-López, J.C.; Crossa, J.; Ramirez, N.L.; Singh, P.; Valladares-Anguiano, F.A. A zero altered Poisson random forest model for genomic-enabled prediction. *G3 Genes Genomes Genet.* **2021**, *11*, jkaa057. [\[CrossRef\]](#)
13. Galli, G.; Sabadin, F.; Yassue, R.M.; Galves, C.; Carvalho, H.F.; Crossa, J.; Montesinos-López, O.A.; Fritsche-Neto, R. Automated Machine Learning: A Case Study of Genomic “Image-Based” Prediction in Maize Hybrids. *Front. Plant Sci.* **2022**, *13*, 845524. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Chen, X.; Chen, D.G.; Zhao, Z.; Zhan, J.; Ji, C.; Chen, J. Artificial image objects for classification of schizophrenia with GWAS-selected SNVs and convolutional neural network. *Patterns* **2021**, *2*, 100303. [\[CrossRef\]](#)
15. Chen, X.; Chen, D.G.; Zhao, Z.; Balko, J.M.; Chen, J. Artificial image objects for classification of breast cancer biomarkers with transcriptome sequencing data and convolutional neural network algorithms. *Breast Cancer Res.* **2021**, *23*, 96. [\[CrossRef\]](#)
16. Khan, A.W.; Garg, V.; Sun, S.; Gupta, S.; Dudchenko, O.; Roorkiwal, M.; Chitkineni, A.; Bayer, P.E.; Shi, C.; Upadhyaya, H.D.; et al. Cicer super-pangenome provides insights into species evolution and agronomic trait loci for crop improvement in chickpea. *Nat. Genet.* **2024**, *56*, 1225–1234. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Bavykina, M.; Kostina, N.; Lee, C.R.; Schafleitner, R.; Bishop-von Wettberg, E.; Nuzhdin, S.V.; Samsonova, M.; Gursky, V.; Kozlov, K. Modeling of Flowering Time in *Vigna radiata* with Artificial Image Objects, Convolutional Neural Network and Random Forest. *Plants* **2022**, *11*, 3327. [\[CrossRef\]](#) [\[PubMed\]](#)
18. Wang, H.; Wang, Z.; Du, M.; Yang, F.; Zhang, Z.; Ding, S.; Mardziel, P.; Hu, X. Score-CAM: Score-Weighted Visual Explanations for Convolutional Neural Networks. In Proceedings of the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), Seattle, WA, USA, 14–19 June 2020; pp. 111–119. [\[CrossRef\]](#)
19. Zhang, X.; Gao, J. Measuring Feature Importance of Convolutional Neural Networks. *IEEE Access* **2020**, *8*, 196062–196074. [\[CrossRef\]](#)
20. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2020**, *128*, 336–359. [\[CrossRef\]](#)
21. Simonyan, K.; Vedaldi, A.; Zisserman, A. Deep Inside Convolutional Networks: Visualising Image Classification Models and Saliency Maps. *arXiv* **2014**, arXiv:1312.6034.
22. Mieth, B.; Kloft, M.; Rodríguez, J.A.; Sonnenburg, S.; Vobruba, R.; Morcillo-Suárez, C.; Farré, X.; Marigorta, U.M.; Fehr, E.; Dickhaus, T.; et al. Combining Multiple Hypothesis Testing with Machine Learning Increases the Statistical Power of Genome-wide Association Studies. *Sci. Rep.* **2016**, *6*, 36671. [\[CrossRef\]](#)
23. Paré, G.; Mao, S.; Deng, W.Q. A machine-learning heuristic to improve gene score prediction of polygenic traits. *Sci. Rep.* **2017**, *7*, 12665. [\[CrossRef\]](#)
24. Nicholls, H.L.; John, C.R.; Watson, D.S.; Munroe, P.B.; Barnes, M.R.; Cabrera, C.P. Reaching the End-Game for GWAS: Machine Learning Approaches for the Prioritization of Complex Disease Loci. *Front. Genet.* **2020**, *11*, 350. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Breiman, L. Random Forests. *Mach. Learn.* **2001**, *45*, 5–32. [\[CrossRef\]](#)
26. Nguyen, T.T.; Huang, J.Z.; Wu, Q.; Nguyen, T.T.; Li, M.J. Genome-wide association data classification and SNPs selection using two-stage quality-based Random Forests. *BMC Genom.* **2015**, *16*, S5. [\[CrossRef\]](#)
27. Szymczak, S.; Holzinger, E.; Dasgupta, A.; Malley, J.D.; Molloy, A.M.; Mills, J.L.; Brody, L.C.; Stambolian, D.; Bailey-Wilson, J.E. r2VIM: A new variable selection method for random forests in genome-wide association studies. *BioData Min.* **2016**, *9*, 7. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Stephan, J.; Stegle, O.; Beyer, A. A random forest approach to capture genetic effects in the presence of population structure. *Nat. Commun.* **2015**, *6*, 7432. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Liu, Y.; Wang, D.; He, F.; Wang, J.; Joshi, T.; Xu, D. Phenotype Prediction and Genome-Wide Association Study Using Deep Convolutional Neural Network of Soybean. *Front. Genet.* **2019**, *10*, 1091. [\[CrossRef\]](#)
30. Chattopadhyay, A.; Sarkar, A.; Howlader, P.; Balasubramanian, V.N. Grad-CAM++: Generalized Gradient-Based Visual Explanations for Deep Convolutional Networks. In Proceedings of the 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 839–847. [\[CrossRef\]](#)
31. Jiang, P.T.; Zhang, C.B.; Hou, Q.; Cheng, M.M.; Wei, Y. LayerCAM: Exploring Hierarchical Class Activation Maps for Localization. *IEEE Trans. Image Process.* **2021**, *30*, 5875–5888. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Sharma, A.; Vans, E.; Shigemizu, D.; Boroevich, K.A.; Tsunoda, T. DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture. *Sci. Rep.* **2019**, *9*, 11399. [\[CrossRef\]](#)
33. Sharma, A.; Lysenko, A.; Boroevich, K.A.; Vans, E.; Tsunoda, T. DeepFeature: Feature selection in nonimage data using convolutional neural network. *Brief. Bioinform.* **2021**, *22*, bbab297. [\[CrossRef\]](#)

34. Sokolkova, A.; Bulyntsev, S.V.; Chang, P.L.; Carrasquilla-Garcia, N.; Igolkina, A.A.; Noujdina, N.V.; von Wettberg, E.; Vishnyakova, M.A.; Cook, D.R.; Nuzhdin, S.V.; et al. Genomic Analysis of Vavilov's Historic Chickpea Landraces Reveals Footprints of Environmental and Human Selection. *Int. J. Mol. Sci.* **2020**, *21*, 3952. [\[CrossRef\]](#) [\[PubMed\]](#)
35. Mairal, J.; Bach, F.; Ponce, J.; Sapiro, G. Online dictionary learning for sparse coding. In Proceedings of the 26th Annual International Conference on Machine Learning, Montreal, QC, Canada, 14–18 June 2009; pp. 689–696. [\[CrossRef\]](#)
36. Bai, Y.; Zhu, Z.; Jiang, G. Blind Quality Assessment of Screen Content Images Via Macro-Micro Modeling of Tensor Domain Dictionary. *IEEE Trans. Multimed.* **2021**, *23*, 13. [\[CrossRef\]](#)
37. Wang, Z.; Yang, J. Diabetic Retinopathy Detection via Deep Convolutional Networks for Discriminative Localization and Visual Explanation. *arXiv* **2019**, arXiv:1703.10757.
38. Xie, W.; Jacobs, C.; Charbonnier, J.P.; Van Ginneken, B. Dense regression activation maps for lesion segmentation in CT scans of COVID-19 patients. *Med. Image Anal.* **2023**, *86*, 102771. [\[CrossRef\]](#) [\[PubMed\]](#)
39. Dunnett, C.W. A Multiple Comparison Procedure for Comparing Several Treatments with a Control. *J. Am. Stat. Assoc.* **1955**, *50*, 1096–1121. [\[CrossRef\]](#)
40. Tong, L. Acetyl-coenzyme A carboxylase: Crucial metabolic enzyme and attractive target for drug discovery. *Cell. Mol. Life Sci.* **2005**, *62*, 1784–1803. [\[CrossRef\]](#)
41. Guna, A.; Volkmar, N.; Christianson, J.C.; Hegde, R.S. The ER membrane protein complex is a transmembrane domain insertase. *Science* **2018**, *359*, 470–473. [\[CrossRef\]](#)
42. Griffié, J.; Shannon, M.; Bromley, C.L.; Boelen, L.; Burn, G.L.; Williamson, D.J.; Heard, N.A.; Cope, A.P.; Owen, D.M.; Rubin-Delanchy, P. A Bayesian cluster analysis method for single-molecule localization microscopy data. *Nat. Protoc.* **2016**, *11*, 2499–2514. [\[CrossRef\]](#)
43. Wiermer, M.; Cheng, Y.T.; Imkampe, J.; Li, M.; Wang, D.; Lipka, V.; Li, X. Putative members of the Arabidopsis Nup107-160 nuclear pore sub-complex contribute to pathogen defense. *Plant J.* **2012**, *70*, 796–808. [\[CrossRef\]](#)
44. Sreekanta, S.; Bethke, G.; Hatsugai, N.; Tsuda, K.; Thao, A.; Wang, L.; Katagiri, F.; Glazebrook, J. The receptor-like cytoplasmic kinase PCRK1 contributes to pattern-triggered immunity against *Pseudomonas syringae* in *Arabidopsis thaliana*. *New Phytol.* **2015**, *207*, 78–90. [\[CrossRef\]](#)
45. Li, Q.; Liu, N.; Liu, Q.; Zheng, X.; Lu, L.; Gao, W.; Liu, Y.; Liu, Y.; Zhang, S.; Wang, Q.; et al. DEAD-box helicases modulate dicing body formation in *Arabidopsis*. *Sci. Adv.* **2021**, *7*, eabc6266. [\[CrossRef\]](#) [\[PubMed\]](#)
46. Tan, Q.K.G.; Irish, V.F. The Arabidopsis Zinc Finger-Homeodomain Genes Encode Proteins with Unique Biochemical Properties That Are Coordinately Expressed during Floral Development. *Plant Physiol.* **2006**, *140*, 1095–1108. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Nole-Wilson, S.; Tranby, T.L.; Krizek, B.A. AINTEGUMENTA-like (AIL) genes are expressed in young tissues and may specify meristematic or division-competent states. *Plant Mol. Biol.* **2005**, *57*, 613–628. [\[CrossRef\]](#) [\[PubMed\]](#)
48. Nour-Eldin, H.H.; Andersen, T.G.; Burow, M.; Madsen, S.R.; Jørgensen, M.E.; Olsen, C.E.; Dreyer, I.; Hedrich, R.; Geiger, D.; Halkier, B.A. NRT/PTR transporters are essential for translocation of glucosinolate defence compounds to seeds. *Nature* **2012**, *488*, 531–534. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Pearce, G.; Yamaguchi, Y.; Barona, G.; Ryan, C.A. A subtilisin-like protein from soybean contains an embedded, cryptic signal that activates defense-related genes. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 14921–14925. [\[CrossRef\]](#)
50. Lim, C.W.; Yang, S.H.; Shin, K.H.; Lee, S.C.; Kim, S.H. The AtLRK10L1.2, Arabidopsis ortholog of wheat LRK10, is involved in ABA-mediated signaling and drought resistance. *Plant Cell Rep.* **2015**, *34*, 447–455. [\[CrossRef\]](#) [\[PubMed\]](#)
51. Lin, R.; Wang, H. Arabidopsis *FHY3/FAR1* Gene Family and Distinct Roles of Its Members in Light Control of Arabidopsis Development. *Plant Physiol.* **2004**, *136*, 4010–4022. [\[CrossRef\]](#)
52. Zheng, Z.; Xia, Q.; Dauk, M.; Shen, W.; Selvaraj, G.; Zou, J. Arabidopsis *AtGPAT1*, a Member of the Membrane-Bound Glycerol-3-Phosphate Acyltransferase Gene Family, Is Essential for Tapetum Differentiation and Male Fertility. *Plant Cell* **2003**, *15*, 1872–1887. [\[CrossRef\]](#)
53. Houben, M.; Van De Poel, B. 1-Aminocyclopropane-1-Carboxylic Acid Oxidase (ACO): The Enzyme That Makes the Plant Hormone Ethylene. *Front. Plant Sci.* **2019**, *10*, 695. [\[CrossRef\]](#)
54. Lisa, M.; Holčápek, M.; Boháč, M. Statistical Evaluation of Triacylglycerol Composition in Plant Oils Based on High-Performance Liquid Chromatography-Atmospheric Pressure Chemical Ionization Mass Spectrometry Data. *J. Agric. Food Chem.* **2009**, *57*, 6888–6898. [\[CrossRef\]](#) [\[PubMed\]](#)
55. Lock, Y.; Snyder, C.L.; Zhu, W.; Siloto, R.M.P.; Weselake, R.J.; Shah, S. Antisense suppression of type 1 diacylglycerol acyltransferase adversely affects plant development in *Brassica napus*. *Physiol. Plant.* **2009**, *137*, 61–71. [\[CrossRef\]](#) [\[PubMed\]](#)
56. Stacey, M.G.; Osawa, H.; Patel, A.; Gassmann, W.; Stacey, G. Expression analyses of Arabidopsis oligopeptide transporters during seed germination, vegetative growth and reproduction. *Planta* **2006**, *223*, 291–305. [\[CrossRef\]](#) [\[PubMed\]](#)
57. Sánchez-García, A.B.; Aguilera, V.; Micol-Ponce, R.; Jover-Gil, S.; Ponce, M.R. Arabidopsis MAS2, an Essential Gene That Encodes a Homolog of Animal NF- κ B Activating Protein, Is Involved in 45S Ribosomal DNA Silencing. *Plant Cell* **2015**, *27*, 1999–2015. [\[CrossRef\]](#)
58. Chen, Q.; Xue, B.; Zhang, M. Genetic Programming for Instance Transfer Learning in Symbolic Regression. *IEEE Trans. Cybern.* **2020**, *52*, 25–38. [\[CrossRef\]](#)
59. Xu, J.; Huang, Y.; Ying, D. Traffic Sign Detection and Recognition Using Multi-Frame Embedding of Video-Log Images. *Remote Sens.* **2023**, *15*, 2959. [\[CrossRef\]](#)

60. Sarita.; Mehrotra, S.; Dimkpa, C.O.; Goyal, V. Survival mechanisms of chickpea (*Cicer arietinum*) under saline conditions. *Plant Physiol. Biochem.* **2023**, *205*, 108168. [[CrossRef](#)]
61. Atieno, J.; Li, Y.; Langridge, P.; Dowling, K.; Brien, C.; Berger, B.; Varshney, R.K.; Sutton, T. Exploring genetic variation for salinity tolerance in chickpea using image-based phenotyping. *Sci. Rep.* **2017**, *7*, 1300. [[CrossRef](#)] [[PubMed](#)]
62. Thompson, C.; Tar'an, B. Genetic characterization of the acetohydroxyacid synthase (AHAS) gene responsible for resistance to imidazolinone in chickpea (*Cicer arietinum* L.). *Theor. Appl. Genet.* **2014**, *127*, 1583–1591. [[CrossRef](#)]
63. Varshney, R.K.; Thudi, M.; Roorkiwal, M.; He, W.; Upadhyaya, H.D.; Yang, W.; Bajaj, P.; Cubry, P.; Rathore, A.; Jian, J.; et al. Resequencing of 429 chickpea accessions from 45 countries provides insights into genome diversity, domestication and agronomic traits. *Nat. Genet.* **2019**, *51*, 857–864. [[CrossRef](#)]
64. Sun, M.; Sun, S.; Jia, Z.; Zhang, H.; Ou, C.; Ma, W.; Wang, J.; Li, M.; Mao, P. Genome-wide analysis and expression profiling of glyoxalase gene families in oat (*Avena sativa*) indicate their responses to abiotic stress during seed germination. *Front. Plant Sci.* **2023**, *14*, 1215084. [[CrossRef](#)] [[PubMed](#)]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.