



Machine Learning Representation of Loss of Eye Regularity in a *Drosophila* Neurodegenerative Model

Sergio Diez-Hermano^{1,2*}, Maria D. Ganfornina¹, Esteban Vegas-Lozano^{3†} and Diego Sanchez^{1†}

¹ Instituto de Biología y Genética Molecular-Departamento de Bioquímica y Biología Molecular y Fisiología, Universidad de Valladolid-CSIC, Valladolid, Spain, ² Departamento de Biodiversidad, Ecología y Evolución, Unidad de Biomatemáticas, Universidad Complutense, Madrid, Spain, ³ Departamento de Genética, Microbiología y Estadística, Universidad de Barcelona, Barcelona, Spain

OPEN ACCESS

Edited by:

Rodolfo Gabriel Gatto,
University of Illinois at Chicago,
United States

Reviewed by:

Santhosh Girirajan,
Pennsylvania State University (PSU),
United States
Praveen Agarwal,
Anand International College
of Engineering, India
Angelique Christine Paulk,
Massachusetts General Hospital,
Harvard Medical School,
United States

*Correspondence:

Sergio Diez-Hermano
serbiodh@gmail.com

† These authors have contributed
equally to this work and share senior
authorship

Specialty section:

This article was submitted to
Brain Imaging Methods,
a section of the journal
Frontiers in Neuroscience

Received: 21 February 2020

Accepted: 27 April 2020

Published: 04 June 2020

Citation:

Diez-Hermano S, Ganfornina MD,
Vegas-Lozano E and Sanchez D
(2020) Machine Learning
Representation of Loss of Eye
Regularity in a *Drosophila*
Neurodegenerative Model.
Front. Neurosci. 14:516.
doi: 10.3389/fnins.2020.00516

The fruit fly compound eye is a premier experimental system for modeling human neurodegenerative diseases. The disruption of the retinal geometry has been historically assessed using time-consuming and poorly reliable techniques such as histology or pseudopupil manual counting. Recent semiautomated quantification approaches rely either on manual region-of-interest delimitation or engineered features to estimate the extent of degeneration. This work presents a fully automated classification pipeline of bright-field images based on orientated gradient descriptors and machine learning techniques. An initial region-of-interest extraction is performed, applying morphological kernels and Euclidean distance-to-centroid thresholding. Image classification algorithms are trained on these regions (support vector machine, decision trees, random forest, and convolutional neural network), and their performance is evaluated on independent, unseen datasets. The combinations of oriented gradient + gaussian kernel Support Vector Machine [0.97 accuracy and 0.98 area under the curve (AUC)] and fine-tuned pre-trained convolutional neural network (0.98 accuracy and 0.99 AUC) yielded the best results overall. The proposed method provides a robust quantification framework that can be generalized to address the loss of regularity in biological patterns similar to the *Drosophila* eye surface and speeds up the processing of large sample batches.

Keywords: *Drosophila melanogaster*, neurodegeneration, rough eye phenotype, spinocerebellar ataxia, machine learning, classification, deep learning

INTRODUCTION

Drosophila melanogaster stands out as one of the key animal models in today's modern genetic studies, with an estimated 75% of human disease genes having orthologs in flies (Reiter et al., 2001). Its growth as a powerful experimental model of choice has been supported by the wide array of genetic and molecular biology tools designed with the fruit fly in mind (Johnston, 2002), easing

Abbreviations: AdaBoost, adaptative boosting; AUC, area under the ROC curve; BN, batch normalization; CNN, convolutional neural network; DT, decision tree; gmr, glass multimer reporter; HOG, histogram of oriented gradients; IREG, regularity index; MLP, multilayer perceptron; NN, neural network; PolyQ, polyglutaminated; RBF, radial basis function; RGB, red, green, blue (colorspace); RF, random forest; ROC, receiver operating characteristic; ROI, region of interest; SCA, spinocerebellar ataxia; SEM, scanning electron micrograph; SGD, stochastic gradient descent; SVM, support vector machine; UAS, upstream activating sequence; WT, wild type.

the creation of genetic deletions, insertions, knock-downs, and transgenic lines. Fly biologists have greatly contributed to our knowledge of mammalian biology, making *Drosophila* the historical premier research system in the fields of epigenetics, cancer molecular networks, neurobiology, and immunology (Wangler et al., 2015). The relative simplicity of *Drosophila* genetics (four pairs of homologous chromosomes, in contrast to 23 in humans) and organization (i.e., $\sim 2 \times 10^5$ neurons in opposition to roughly 10^{11} neurons in humans) makes the fruit fly an especially well-suited model for the analysis of subsets of phenotypes associated with complex disorders.

Specifically, the retinal system in *Drosophila* has been widely used as an experimental setting for high-throughput genetic screening and for testing molecular interactions (Thomas and Wassarman, 1999). Eye development is a milestone in the *Drosophila* life cycle, with a massive two-thirds of the essential genes in the fly genome required at some point during the process (Thaker and Kankel, 1992; Treisman, 2013). Therefore, it constitutes an excellent playground to study the genetics underlying general biological phenomena, from the basic cellular and molecular functions to the pathogenic mechanisms involved in multifactorial human diseases, such as diabetes or neurodegeneration (Garcia-Lopez et al., 2011; Lenz et al., 2013; He et al., 2014).

The fruit fly compound eye is a biological system structured as a stereotypic array of 800 simple units, called ommatidia, which display a highly regular hexagonal pattern (Figure 1). This strict organization precisely allows to evaluate the impact of altered gene expression and mutated proteins on the external eye morphology and to detect subtle alterations on the ommatidia geometry due to cell degeneration. One special type of cellular deterioration largely studied using *Drosophila* retina encompasses polyglutamine-based neurodegenerative diseases, namely, Huntington's and spinocerebellar ataxias (SCA) (Ambegaokar et al., 2010).

The overexpression of polyQ-expanded proteins *via* the UAS/Gal4 system in the fly retina results in a depigmented, rough eye phenotype caused by the loss of interommatidial bristles (see the wild-type pattern in the inset of Figure 1B), ommatidial fusion, and necrotic tissue (Figure 2). The vast

majority of studies assessing the rough eye morphology rely on qualitative examination (i.e., visual inspection) of its external appearance to manually rank and categorize mutations based on their severity (Roederer et al., 2005; Bilen and Bonini, 2007; Cukier et al., 2008). Even though evident degenerated phenotypes are easily recognizable, weak modifiers or subtle alterations may go undetected for the naked eye. Quantitative approaches addressing this issue involve histological preparations from which to evaluate the retinal thickness and the regularity of the hexagonal array or scoring scales for the presence of expected features in the retinal surface (Jonshon and Cagan, 2009; Jenny, 2011; Caudron et al., 2013; Mishra and Knust, 2013; Song et al., 2013). Recently, there have been efforts to fully computerize the analysis of *Drosophila*'s rough eye phenotype in bright-field and scanning electron micrograph (SEM) images in the form of ImageJ plugins, called FLEYE and Flynotiper (Diez-Hermano et al., 2015; Iyer et al., 2016). Whereas both methods propose automatized workflows, the former prompts the user to manually delimit the region of interest (ROI) to extract the hand-crafted features from it, which serve as input to a statistical model and finally output a regularity index (IREG) to the user. The second method relies upon a single engineered feature and lacks statistical background to support it.

Hence, there is a need to tackle a fully automatized, statistically multivariate assessment of *Drosophila* eye's quantification, given its utmost relevance as a simple, yet comprehensive, model for testing general biology hypotheses and human neurological diseases. Particularly, machine learning algorithms have proven to be incredibly efficient image classifiers during the past decade (Bishop, 2006), rapidly permeating in the fields of cell biology and biomedical image-based screening (Sommer and Gerlich, 2013; Chessel, 2017; Tyagi, 2019). Machine learning methods greatly ease the analysis of complex multi-dimensional data by learning processing rules from examples that can be later on generalized to classify new, unseen data (Figure 3A).

The machine learning techniques typically applied to image classification includes support vector machines (SVM) (Ben-Hur et al., 2008; Chauhan et al., 2019), decision trees (DT) (Orrite et al., 2009), random forests (RF) (Schroff et al., 2008), and neural networks (NNs) (Giacinto and Roli, 2001). Alongside processing power and graphic card-dedicated coding, deep learning methods have exponentially grown in importance during the last few years (LeCun et al., 2015; Po-Hsien et al., 2015). The conventional machine learning algorithms aforementioned require data processing and feature enrichment prior to the training phase as they are not suited to work with raw input. In contrast, the deep learning procedures are general-purpose learners in the sense that they can be fed with raw data, automatically suppress irrelevant information, and select discriminant characteristics, composing simple layers of non-linear transformations into a higher, more abstract level of representation (Figure 3B). The convolutional neural networks (CNNs) are a well-known architecture for deep learning and have been continuously outperforming the previous machine learning techniques, especially in computer vision and audio recognition (Po-Hsien et al., 2015). With the increasing availability of large biological datasets, its popularity in bioinformatics and bioimaging has quickly escalated, and

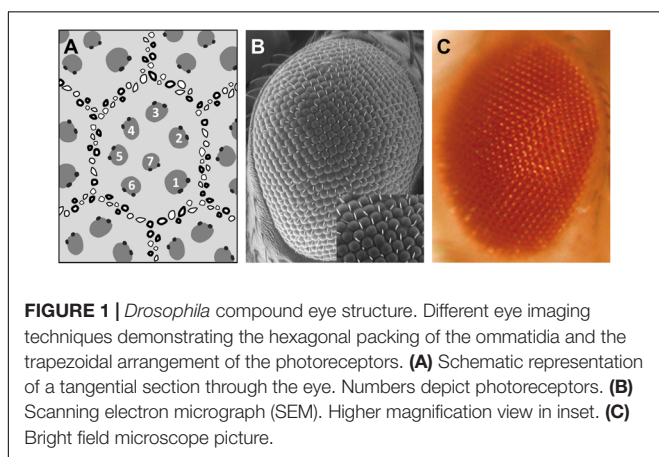


FIGURE 1 | *Drosophila* compound eye structure. Different eye imaging techniques demonstrating the hexagonal packing of the ommatidia and the trapezoidal arrangement of the photoreceptors. (A) Schematic representation of a tangential section through the eye. Numbers depict photoreceptors. (B) Scanning electron micrograph (SEM). Higher magnification view in inset. (C) Bright field microscope picture.

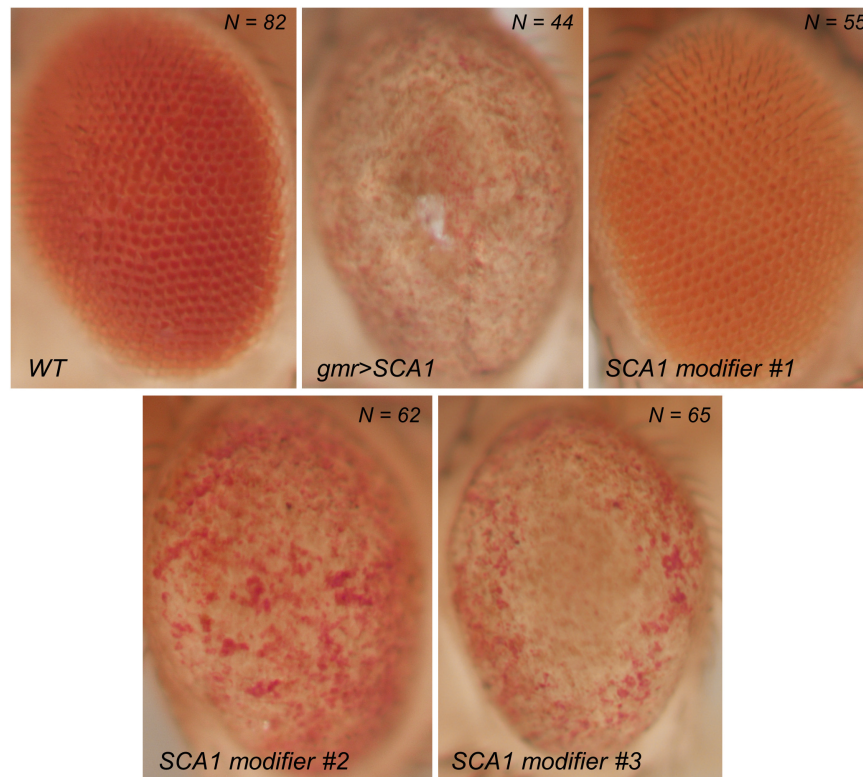


FIGURE 2 | Bright field photographs of rough eye phenotype surfaces. SCA1 gene modifiers can be tested on the fly eye using the UAS/Gal4 system. Complete loss of surface regularity and depigmentation can be appreciated between the WT and SCA1 phenotypes. SCA1 modifiers show intermediate levels of degeneration.

currently CNNs are addressing problems hardly resolvable by former top-notch analysis techniques (Angermueller et al., 2016; Chen et al., 2016; Kraus et al., 2016; Spanhol et al., 2016; Yang et al., 2016; Anwar et al., 2018; Badar et al., 2020). The striking advantage of these networks is that a feature's hand-crafting and engineering are completely avoided as they implement functions insensitive to perturbations, thanks to the multilayer mapping representation of discriminant details.

The novelty of the present work consists in applying and comparing the different image classification strategies mentioned so far in an extensively used biological model, *D. melanogaster*, which has been scarcely addressed before and is in dire need of a state-of-the-art quantification framework.

MATERIALS AND METHODS

Fly Lines and Maintenance

All stocks and crosses were grown in a temperature-controlled incubator at 25°C, 60% relative humidity, and under a 12-h light-dark cycle. They were fed on conventional medium containing wet yeast 84 g/L, NaCl 3.3 g/L, agar 10 g/L, wheat flour 42 g/L, apple juice 167 ml/L, and propionic acid 5 ml/L. To drive transgene expression to the eye photoreceptor, we used the line *gmr:GAL4*. Rough eye phenotype was triggered using the UAS:hATXN182Q transgene (Fernandez-Funez et al., 2000)

that models human type 1 spinocerebellar ataxia (SCA1), and different UAS:modifier-gene constructs were used to test the system capability to recognize intermediate phenotypes.

Sample Size

A total of 308 image files were saved using NIS-Elements software in TIFF format. The number of pictures by category is as follows: 82 wild type (WT), 44 *gmr > SCA1*, 55 modifier #1, 62 modifier #2, and 65 modifier #3.

External Eye Surface Digital Imaging

Digital pictures (2,880 × 2,048 pixels) of the surface of fly eyes were taken with a Nikon DS-Fi3 digital camera and viewed with a Nikon SMZ1000 stereomicroscope equipped with a Plan Apo× 1 WD70 objective. The flies were anesthetized with CO₂ and their bodies were immobilized on dual adhesive tape, with their heads oriented to have an eye parallel to the microscope objective. The fly eyes were illuminated with a homogeneous fiber optic light passing through a translucent cylinder so that the light rays were dispersed and did not directly reach the eyes. The images taken with this method show a better representation of the surface retinal texture in contrast to the pictures where light fell upon the eye and the lens' reflection was captured by the camera, forming bright-spotted grids. The additional settings include an 8× optical zoom in the stereomicroscope.

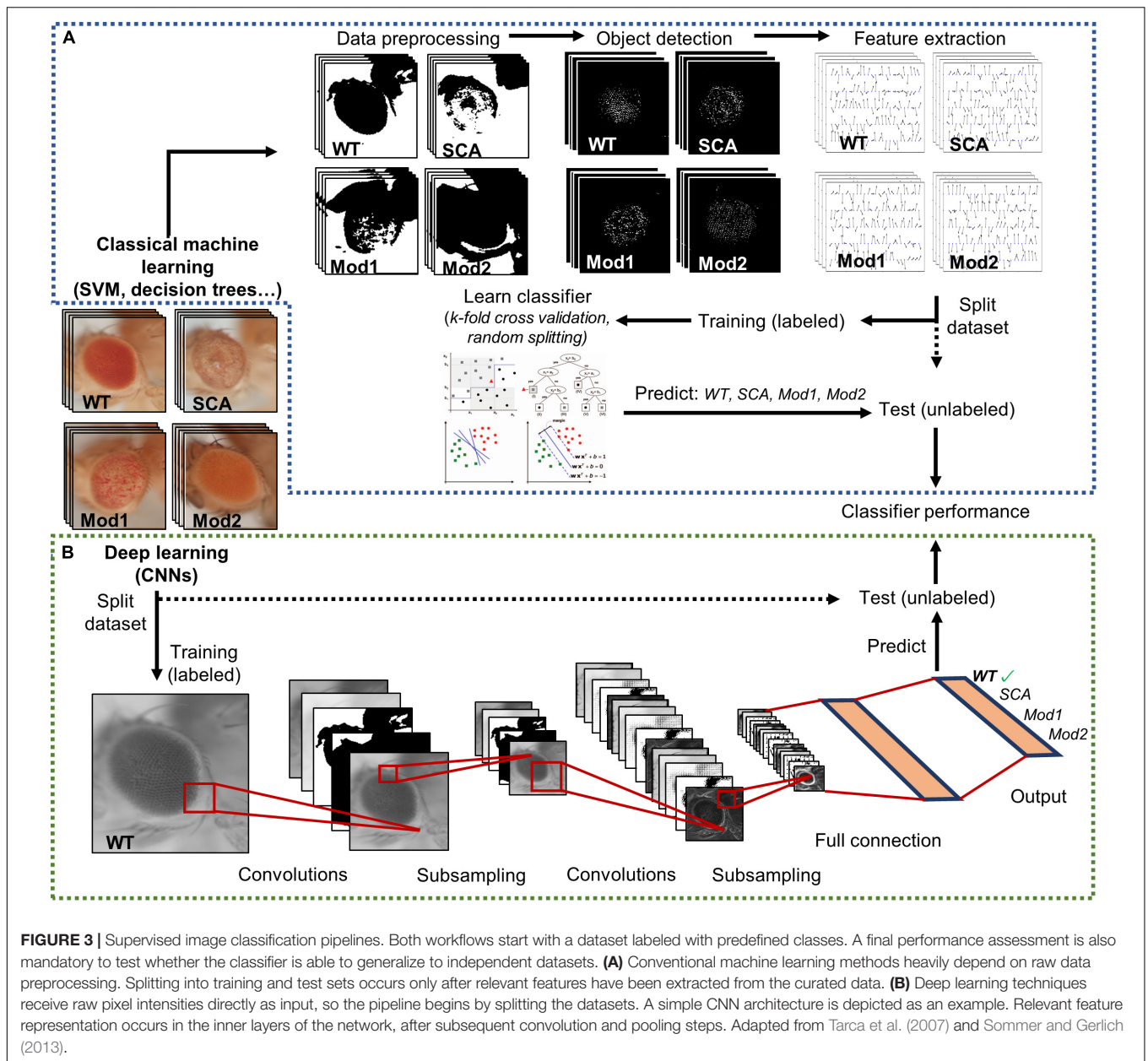


FIGURE 3 | Supervised image classification pipelines. Both workflows start with a dataset labeled with predefined classes. A final performance assessment is also mandatory to test whether the classifier is able to generalize to independent datasets. **(A)** Conventional machine learning methods heavily depend on raw data preprocessing. Splitting into training and test sets occurs only after relevant features have been extracted from the curated data. **(B)** Deep learning techniques receive raw pixel intensities directly as input, so the pipeline begins by splitting the datasets. A simple CNN architecture is depicted as an example. Relevant feature representation occurs in the inner layers of the network, after subsequent convolution and pooling steps. Adapted from Tarca et al. (2007) and Sommer and Gerlich (2013).

ROI Selection Algorithm

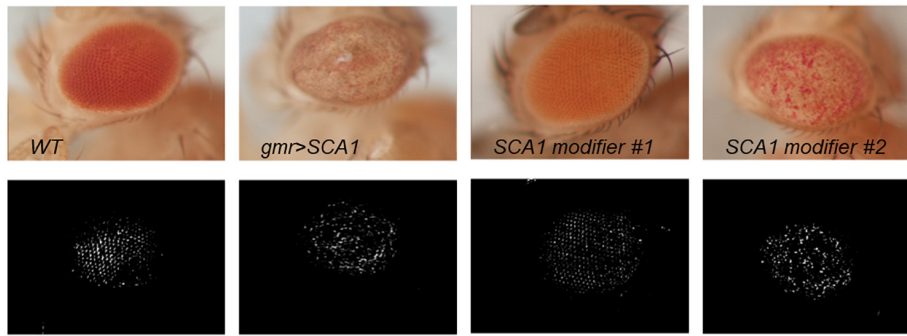
All image analyses were performed using R programming language (R Core Team, 2018). The eye images in red/green/blue (RGB) color space were first resized to one-fourth of their original resolution to help fit the image data to the memory capacity of the computer system used. White TopHat morphological transformation with a disc kernel of size 9 was applied using the package EBImage (Pau et al., 2010). The transformed images are converted to grayscale and thresholded to keep only pixels with intensity >0.99 quantile. The overall centroid of the remaining pixels is estimated using the Weiszfeld L1-median (Vardi and Cun-Hui, 2000). For each pixel, the Euclidean distance to the centroid is calculated, and those with distances >0.8 quantile are discarded. A 0.90 confidence

level ellipse is estimated on the final selected pixels, and its area is superimposed to the original resized picture to extract the final ROI.

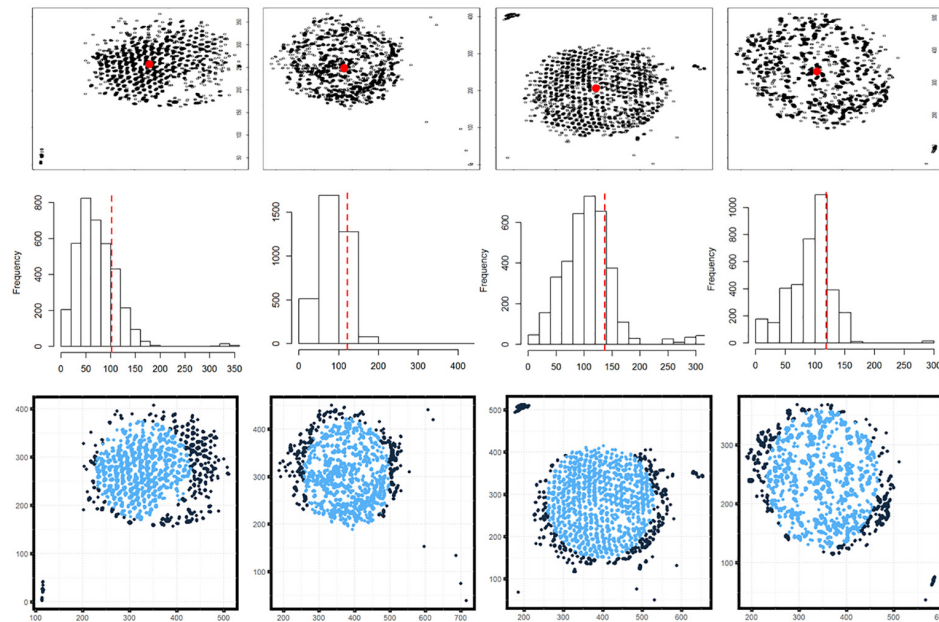
HOG Descriptor and Machine Learning Classifiers

Firstly, RGB ROIs were converted to grayscale while maintaining the original luminance intensities. The histogram of gradient (HOG) features was extracted using the OpenImageR package (Mouselimis, 2017). A 5×5 cell descriptor with five orientations covering a gradient range of $0\text{--}180^\circ$ was estimated per cell in the gradient, resulting in a final 125-dimensional vector for each ROI.

A White TopHat transformation



B Distances to centroid and quantile selection



C Confidence ellipse and ROI selection

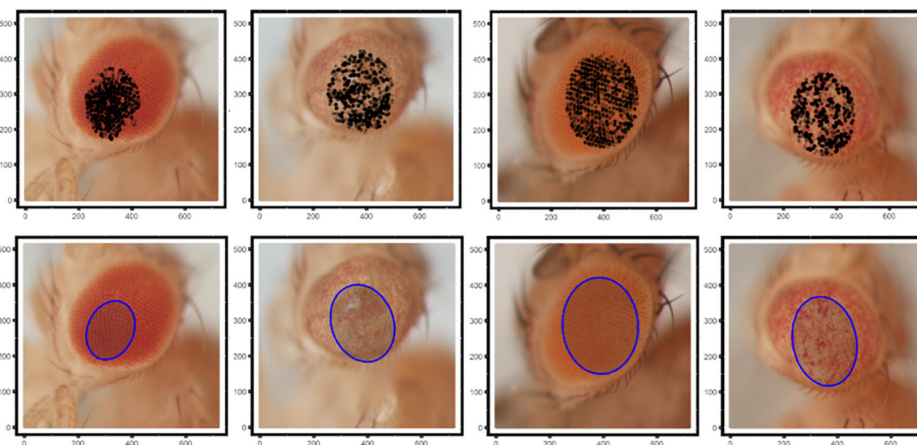


FIGURE 4 | *Drosophila* eye ROI detection strategy. Representative examples of healthy and degenerated eyes are shown. **(A)** Morphological transformation and intensity thresholding extract pixels mostly contained within the eye. **(B)** Euclidean distance to the centroid (red dot) and frequency histogram for quantile selection. Dark blue points are discarded as potential pixel outliers outside the eye limit. **(C)** Selected pixels are superposed to the original image and those within the area of a 0.90 confidence ellipse are extracted as the final ROI (blue shaded ellipse).

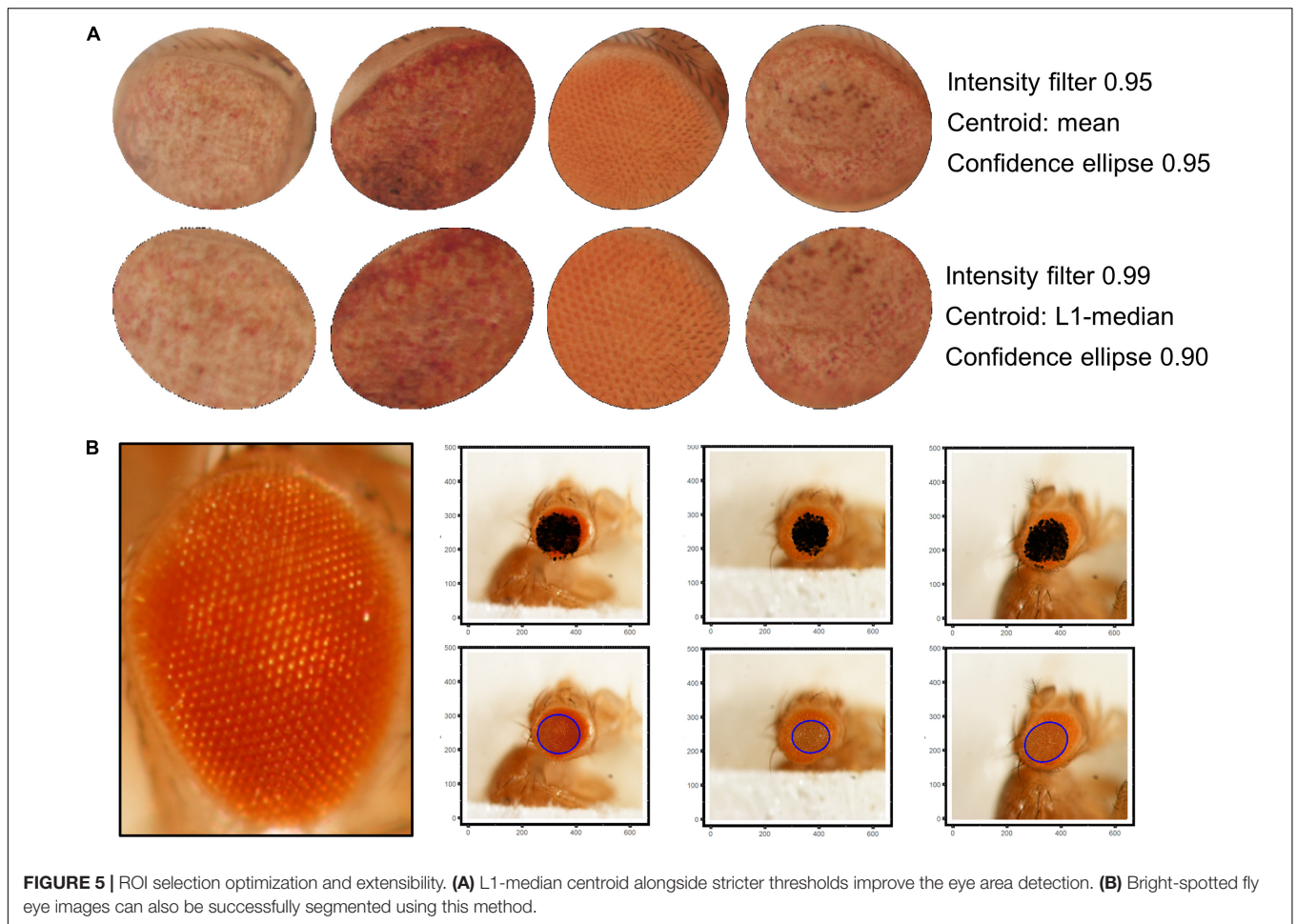


TABLE 1 | Machine learning classifier confusion matrix.

Predicted	Reference																			
	WT				gmr > SCA				SCA modifier #1				SCA modifier #2				SCA modifier #3			
WT	20	16	20	20	0	2	2	2	1	4	3	1	0	0	0	1	0	4	3	1
gmr > SCA	0	2	0	0	11	7	7	8	0	2	0	0	0	1	1	0	0	2	3	0
SCA modifier #1	0	0	0	0	0	0	2	0	12	7	10	11	0	0	0	0	0	0	0	0
SCA modifier #2	0	1	0	0	0	0	0	0	0	0	0	1	14	12	11	12	0	3	0	0
SCA modifier #3	0	1	0	0	0	2	0	1	0	0	0	0	1	2	3	2	16	7	10	15

Colors scheme: SVM, DT, Boost DT, RF.

The SVM, DT, and RF algorithms were trained on the extracted HOG features. The dataset was split into training and test sets with a 75/25 ratio using stratified random sampling to ensure class representation. The modeling strategy for all classifiers included cross-validation to assess generalization, grid search for parameter selection and performance evaluation on test set *via* confusion matrix, global accuracy, Kappa statistic, and multiclass pairwise area under the curve (AUC) (Ferri et al., 2003). We tested a radial basis function (RBF) kernel SVM, DT, adaptative boosting DT, and 1,000-trees RF using the R packages *kernelab*, *C50*, and *caret* (Karatzoglou et al., 2004; Kuhn et al., 2015).

Deep Learning Classifiers

The extracted ROIs were resized to a 224 × 224 × 3 RGB array and stored in vectorized form, resulting in a final data frame of 308 × 150,528 dimensions. The dataset was split into training and test sets with a 75/25 ratio using stratified random sampling to ensure class representation. We further confirmed that the training and test partitions were representative of the sample variability *via* a loss plot (**Supplementary Figure S1**). Two CNNs were trained on this data:

- (i) A simple CNN trained from scratch, with hyperbolic tangent as activation function, two convolutional layers,

TABLE 2 | Machine learning performance evaluation metrics on test data.

Metric	Classifier			
	SVM RBF	DT	AdaBoost DT	1000 RF
Accuracy	0.973 (0.907–0.997)	0.653 (0.535–0.760)	0.773 (0.662–0.862)	0.880 (0.784–0.944)
Kappa	0.966	0.560	0.711	0.847
Multiclass AUC	0.978	0.665	0.763	0.906

True positives are shaded in gray.

two pooling layers, two fully connected layers (200 and five nodes), 30 epochs, and a typical softmax output. Each convolutional layer uses a 5×5 kernel and 20 or 50 filters, respectively. The pooling layers apply a classical “max pooling” approach. All the parameters in kernels, bias terms, and weight vectors are automatically learned by back-propagation with a learning rate equal to 0.05 and a stochastic gradient descent (SGD) optimizer to ensure that the magnitude of the updates stayed small (Bottou, 2010).

- (ii) A fine-tuned CNN using an ImageNet pretrained model with a batch-normalization network structure (Deng et al., 2009; Ioffe and Szegedy, 2015), 30 epochs, a very slow learning rate (0.05), and a SGD optimizer. The final fully connected (five nodes) and softmax output layers are tuned to fit the new fly eye ROIs.

For the CNN training, the R package *MXNet* compiled for the central processing unit (CPU) was used (Tianqi et al., 2015). Performance was assessed in terms of confusion matrix and global accuracy using the *caret* package (Kuhn, 2008).

RESULTS

Automatized Detection of *Drosophila* Eye ROIs From Bright-Field Images

The first step in the quantification workflow is the extraction of pixels corresponding to the fly eye from the rest of the image. One concern is that the eye is not flat but convex in morphology, so under white light only the central surface is at the camera focus. To address this issue, white TopHat morphological transformations were performed, defined as the difference between the input image and its opening by a structuring kernel. The opening operation involves erosion followed by a dilation of the image, retrieving the objects of the input image that are simultaneously smaller than the structuring element and brighter than their neighbors.

Best results were obtained using a 9×9 disc-shaped kernel followed by a thresholding of pixels with intensities over the 0.99 percentile (Figure 4A). Afterward, the centroid of the selected pixels was calculated as the L1-median, which is a more robust estimator of the central coordinates than the arithmetic mean. Points with Euclidean distance to the centroid greater than 0.8 percentile are more likely to lie outside the eye area and were discarded (Figure 4B). A 0.90 confidence ellipse calculated on the selected pixels conforms the area of the final ROI, which

was superimposed and cropped from the original eye image (Figure 4C). As can be appreciated in the example images, the method is invariant to the location of the eye within the image. Various combinations of the thresholds and the centroid estimator were tested (Figure 5A). The proposed segmentation method also works well on bright-field images where light falls directly onto the ommatidium and the eye is seen as a region enriched in reflection spots (Figure 5B). The full array of the final ROIs is represented in Supplementary Figure S2.

We also addressed whether there was any anatomical preference in the ROI extraction that could be biasing the classification procedure. To this extent, heatmaps overlaying all the elliptic ROIs were generated genotype-wise (Supplementary Figure S3). There were a few instances in which the selected ROI included areas outside the eye. This happened more easily in the modifier genotypes and can be seen in the figure as shades that lay outside the eye border. Overall, the ROI extraction seems to be robust against the different ommatidial distribution and eye shapes.

HOG Feature Extraction

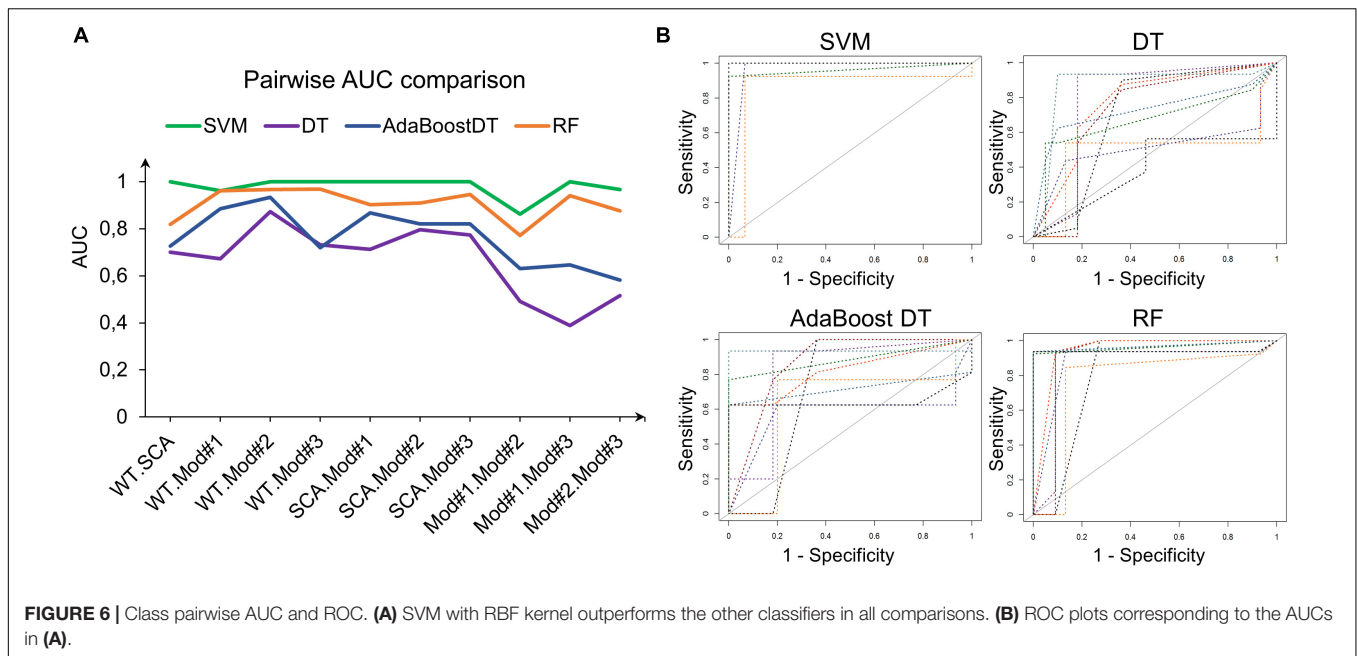
Region of interests cannot be directly input to classical machine learning techniques, so the information contained in their pixels must be extracted beforehand. This is done by estimating a HOG, which can be interpreted as a feature descriptor of a picture that outputs summarized information about predominant shapes and structures. The HOG technique starts by dividing the picture into cells and identifying whether a given cell is an edge or not. HOG provides the edge direction as well, which is done by extracting the gradient and orientation (magnitude and direction) of the edges across neighbor cells. These cells comprise the local regions of related pixels, from which the HOG generates a histogram using the gradients and the orientations of the pixel values, hence the name “histogram of oriented gradients.”

Prior to the HOG extraction, the ROIs were transformed to grayscale, preserving the luminance of the original RGB image. Then, a 125-dimensional feature vector is extracted for each ROI, representing the frequency of a certain gradient within the image (Supplementary Figure S4). The matrix formed by the 125-D vectors of all ROIs conforms to the input for the machine learning classifiers.

Note that HOG is only used to feed the classical machine learning algorithms (SVM, DT, and RF), not the deep learning CNN, which directly uses the pixels' values as input. This is due to the internal structure of the CNN, the inner layers of which serve as border and edge detectors themselves. This is a reason that led us to believe that pigmentation was not affecting the classification procedure, given that all the methods we used relied on structure detectors rather than color differences.

Comparison of Machine Learning Classifiers

Support vector machine with RBF kernel, DT, AdaBoost DT, and 1,000-trees RF algorithms were tested on the extracted HOG features. The sample consisted in 308 fly eye images distributed in five different phenotype classes with varying



degrees of retinal surface degeneration. The data were split using stratified random sampling in 75% training and 25% test set. The optimal parameters for each classifier were found using 10-fold cross-validation on the training set. **Table 1** shows the confusion matrix, and **Table 2** represents the global accuracy, Kappa statistic, and multiclass AUC, defined as the average AUC of class pairwise comparisons (**Figure 6A**), which was calculated on the test data. The pairwise receiver operating characteristic (ROC) plots are represented in **Figure 6B**.

In general, the four classifiers performed fairly well on unseen data. Both DT algorithms fell on the low spectrum either in accuracy and AUC (<0.80), whereas RF achieved a remarkable AUC of 0.90. Overall, SVM accomplished the best results among

all the error metrics evaluated, with a global accuracy of 0.97 (0.90–0.99), Kappa of 0.96, and a multiclass AUC of roughly 0.98. The parameters that yield these results were a Gaussian kernel, a cost penalty = 1, and sigma = 0.005. The WT eyes were the most correctly classified phenotype by the four methods.

From the SVM estimated class probabilities, it is possible to derive a IREG that ranges from 0 (total degeneration) to 1 (healthy eye) (Diez-Hernando et al., 2015). It is based on the knowledge of the degeneration intensity of the phenotypes involved in the model: WT < modifier #1 < modifier #2 < modifier #3 < SCA1, from absence to full presence of rough eye phenotype. IREG is then calculated as:

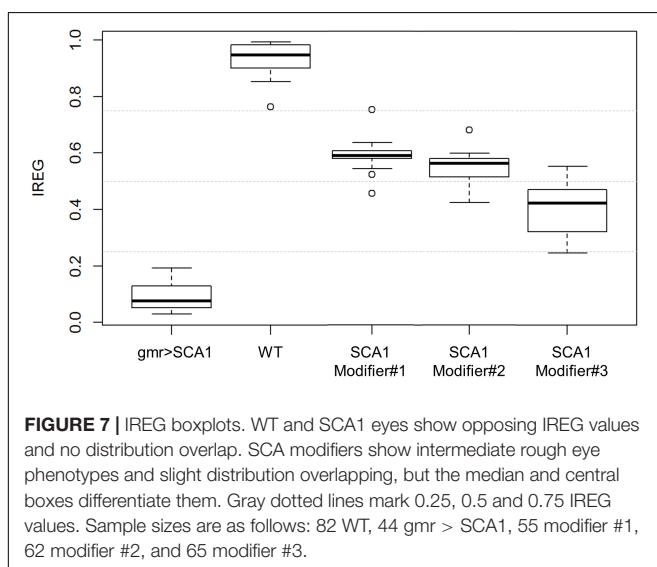
$$\text{IREG} = \frac{4 \cdot P(\text{eye} = \text{WT}) + 3 \cdot P(\text{eye} = \text{Mod}\#1) + 2 \cdot P(\text{eye} = \text{Mod}\#2) + P(\text{eye} = \text{Mod}\#3)}{4}$$

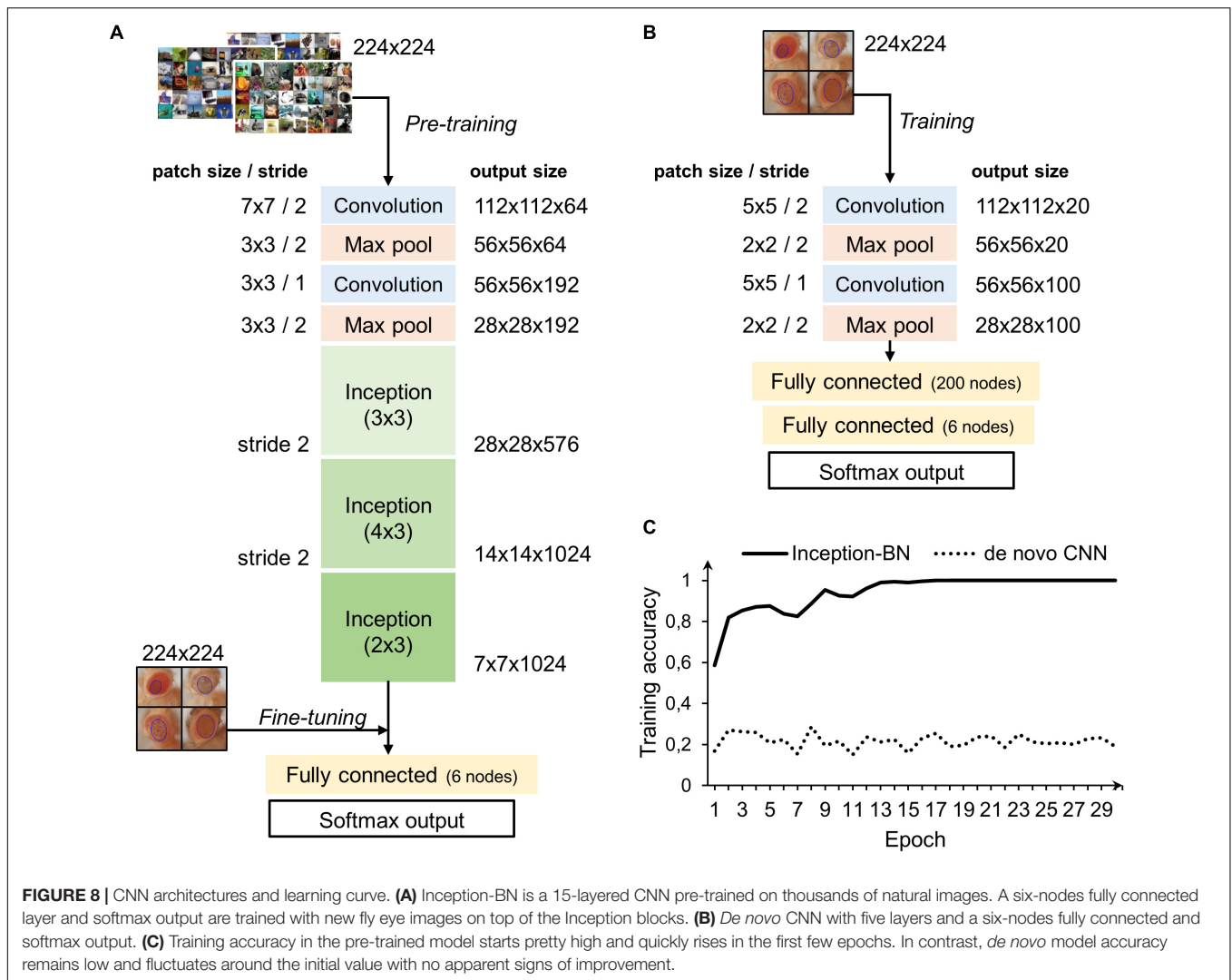
when estimated on the test data, the IREG distribution fits to the expected values and properly reflects the intrinsic variability of the fly model and the rough eye phenotype (**Figure 7**).

Deep Learning Classifiers

In contrast with the previous machine learning classifiers that needed a transformation of the cropped images into an enriched feature space (HOG), deep learning algorithms directly use the ROI pixel intensity arrays as input. The features are automatically learned during the learning process, from gross edge and contour detection to discrimination of fine details the deeper the layer in the network.

Two different strategies were followed to train the deep networks: learning a *de novo* model and transfer learning. The latter approach takes advantage of CNNs pre-trained on very large samples, which is especially well suited for classifying new small datasets as the majority of patterns and motifs commonly



**TABLE 3 |** Convolutional neural networks classifier confusion matrix.

Predicted	Reference									
	WT		gmr > SCA		SCA modifier #1		SCA modifier #2		SCA modifier #3	
WT	20	20	0	11	0	13	0	15	0	16
gmr > SCA	0	0	11	0	0	0	0	0	0	0
SCA modifier #1	0	0	0	0	13	0	0	0	0	0
SCA modifier #2	0	0	0	0	0	0	14	0	0	0
SCA modifier #3	0	0	0	0	0	0	1	0	16	0

Color scheme: Inception-BN *de novo* CNN.

found in the images are already known to the model internal representation. Thus, it is only necessary to fine-tune the final layers to learn the particularities of the new images, which is many times faster than training a CNN from scratch and does not require thousands of labeled examples. The architectures of both *de novo* and pre-trained CNN are depicted in **Figures 8A,B**. The pre-trained model chosen uses the inception structure, characterized by including mini-batch normalization (BN) for

each training epoch, which allows for high learning rates and acts as regularizer. In comparison, the *de novo* CNN is much shallower due to computational constraints.

Accuracy during the training phase is usually a reliable indicator of a CNN capability to learn the discriminative features with the available sample size (**Figure 8C**). The curve of the *de novo* CNN is a clear sign that either the network is not deep enough or the training sample is too small for the complexity

TABLE 4 | Convolutional neural networks performance evaluation metrics on test data.

Metric	Classifier	
	Inception-BN	<i>De novo</i> CNN
Accuracy	0.986	0.146
Multiclass AUC	0.997	0.5

True positives are shaded in gray.

of the classification task at hand. One major concern with the pre-trained inception-BN was the possibility that the network was memorizing the training set, given the few epochs it needed to achieve perfect training accuracy. A performance assessment in an independent test set of unseen images gave impressive accuracy and AUC values close to 1 (Tables 3, 4), refuting the possibility of overfitting. The CNN trained from scratch predicted every eye to be WT, indicative of the weak classifying rule learned in training.

Transfer learning CNN was also the only model capable of properly quantifying the pictures taken under drastically different illumination techniques, relative to the sample that it was trained on (dispersed indirect light). The genetic background also differed from the training sample and corresponds to tau-related neurodegeneration, the accumulation of which contributes to the pathology of Alzheimer disease (Lasagna-Reeves et al., 2016; Rousseaux et al., 2016; Galasso et al., 2017). **Supplementary Figure S5** shows how IREG estimations coming from CNN were more representative of the surface regularity than IREG coming from the SVM, which was the top-performance machine learning technique. In fact, the SVM predicted all eyes to have essentially the same IREG value. Additionally, retraining all models with our original sample of bright-spotted eyes (**Supplementary Figure S2B**) resulted in transfer learning CNN being the only model that could successfully generalize to a set of new eyes (**Supplementary Table S1**).

Given its classification accuracy and versatility, transfer learning with the pre-trained inception-BN model is arguably the top performer classifier among all the methods tested in this work.

DISCUSSION

The present work provides a novel and fully automated method to quantitatively assess the degeneration intensity of the fruit flies' compound eye using reliable and robust state-of-the-art machine learning techniques. This new method consists in the acquisition of bright-field images from the external retinal surface, the automatic extraction of a ROI enriched in information of the eye morphology, and a classification algorithm built around a pre-trained deep learning algorithm, fine-tuned to the particularities of the eye degeneration's images. Additionally, a model based on the combination of HOG features extraction and Gaussian kernel SVM offered performance on par with the CNN and, in fact, required much less training time.

In contrast with previous quantification approaches (Caudron et al., 2013; Diez-Hermano et al., 2015; Iyer et al., 2016), this method does not rely on patterns that are created by light reflecting in the eye lenses, so it can be applied to extract ROIs from a variety of illumination conditions. While performing the experiment to validate the method, we have estimated the total time that it takes for a researcher to analyze an experimental group of 50 flies: 1.5 h from anesthetizing the flies until the final IREG plot was statistically assessed. It is noteworthy that the most part of that estimation was devoted to capturing the images, which is a mandatory step, whether the flies are to be manually or automatically classified later on. The proposed pipeline can process a $2,880 \times 2,048$ -resolution image in less than 10 s and batches of 50 images in approximately 90 s, depending on the hardware that it runs on.

One of the major goals of this work was to analyze the potentiality of deep learning techniques to extract feature maps directly from the raw pixel array, which could be used as input to other conventional machine learning algorithms (i.e., SVM). Due to computational constraints, it was not possible to tune up the graphics processing unit-compiled versions of the software utilized, and the prohibitive CPU computational time and memory usage in its absence made the evaluation of the former objective not feasible. HOG was chosen as an alternative descriptor, given its successful application in object detection (Dalal and Triggs, 2005; Orrite et al., 2009; Li et al., 2012), and ended up resulting in a surprisingly powerful classifier in combination with conventional SVM. Nonetheless, the pre-trained CNN will still be preferable for pictures taken under illumination conditions different from the ones the models were trained on as it has been shown to have greater discriminative power.

A drawback of CNNs is the staggering amount of labeled training examples that they need to learn adequate internal representations of image patterns and motifs. Although the sample size in *Drosophila* experiments ranks among the largest of any animal model in genetics, it is still a titanic effort to go beyond 1,000 images in a typical fruit fly assay. This limitation affected the performance of the *de novo* CNN, which led to the alternative strategy of transfer learning. Using inception-BN, a CNN pre-trained on millions of natural images (Ioffe and Szegedy, 2015), proved to be a well-thought solution that definitely opens up the field of deep learning to small-scale biology setups.

Future lines of work include developing the fly eye detection algorithm further to make it extensible to other image capturing techniques (i.e., SEM). A more immediate priority is the creation of a user-friendly Shiny application (Winston et al., 2017) that will allow the researcher to tweak the ROI selection parameters to fit the peculiarities of its own dataset prior to the degeneration quantification. Depending on the particular hardware settings, the app may also offer the user the possibility to train its own SVM or deep learning model.

The aim of this work is to provide a workflow that results in a quantitative assessment of the degree of eye degeneration of hundreds of flies in a quick and unbiased manner. This makes our method particularly suitable for discriminating potential genetic rescues or aberrations. We believe that our algorithm

could be easily implemented in fully robotized environments as the final quantification step. The highlighted strengths of the proposed framework will enhance the sensitivity of high-throughput genetic screens based on rough eye phenotypes and demonstrate that fly eye imaging is a top-notch technique for the quantitative modeling of human diseases.

DATA AVAILABILITY STATEMENT

The datasets generated and analyzed for this study, alongside the R scripts, can be found in the FigShare repository at <https://tinyurl.com/un8tacu>.

AUTHOR CONTRIBUTIONS

SD-H performed the experiments, took the fly eye pictures, analyzed the data, and wrote the manuscript. MG and DS designed the genetic background of the flies, took the fly eye pictures, and helped with the manuscript writing. EV-L analyzed the data and helped with the manuscript writing.

REFERENCES

- Ambegaokar, S. S., Roy, B., and Jackson, G. R. (2010). Neurodegenerative models in *Drosophila*: polyglutamine disorders, Parkinson disease, and amyotrophic lateral sclerosis. *Neurobiol. Dis.* 40:1. doi: 10.1016/j.nbd.2010.05.026
- Angermueller, C., Parnamaa, T., Parts, L., and Stegle, O. (2016). Deep learning for computational biology. *Mol. Syst. Biol.* 12:7. doi: 10.15252/msb.20156651
- Anwar, S. M., Majid, M., and Qayyum, A. (2018). Medical image analysis using convolutional neural networks: a review. *J. Med. Syst.* 42:226. doi: 10.1007/s10916-018-1088-1
- Badar, M., Haris, M., and Fatima, A. (2020). Application of deep learning for retinal image analysis: a review. *Comput. Sci. Rev.* 35:100203. doi: 10.1016/j.cosrev.2019.100203
- Ben-Hur, A., Ong, C. S., Sonnenburg, S., Schölkopf, B., and Rätsch, G. (2008). Support vector machines and kernels for computational biology. *PLoS Comput. Biol.* 4:173. doi: 10.1371/journal.pcbi.1000173
- Bilen, J., and Bonini, N. M. (2007). Genome-wide screen for modifiers of ataxin-3 neurodegeneration in *Drosophila*. *PLoS Genet.* 3:30177. doi: 10.1371/journal.pgen.0030177
- Bishop, C. M. (2006). *Pattern Recognition And Machine Learning*. Berlin: Springer.
- Bottou, L. (2010). "Large-scale machine learning with stochastic gradient descent," in *Proceedings of the COMPSTAT'2010*, eds Y. Lechevallier, and G. Saporta, (Cham: Physica-Verlag HD).
- Caudron, Q., Lyn-Adams, C., Aston, J. A. D., Frenguelli, B. G., and Moffat, K. G. (2013). Quantitative assessment of ommatidial distortion in *Drosophila melanogaster*: a tool to investigate genetic interactions. *J. Neurogenet.* 24:87.
- Chauhan, V. K., Dahiya, K., and Sharma, A. (2019). Problem formulations and solvers in linear SVM: a review. *Artif. Intell. Rev.* 52, 803–855. doi: 10.1007/s10462-018-9614-6
- Chen, C. L., Mahjoubfar, A., Tai, L. C., Blaby, I. K., Huang, A., Niazi, K. R., et al. (2016). Deep learning in label-free cell classification. *Sci. Rep.* 6:21471. doi: 10.1038/srep21471
- Chessel, A. (2017). An Overview of data science uses in bioimage informatics. *Methods* 115, 110–118. doi: 10.1016/j.ymeth.2016.12.014
- Cukier, H. N., Perez, A. M., Collins, A. L., Zhou, Z., Zoghbi, H. Y., and Botas, J. (2008). Genetic modifiers of MeCP2 function in *Drosophila*. *PLoS Genet.* 4:179. doi: 10.1371/journal.pgen.1000179

FUNDING

This work was supported by grants to MG and DS [Ministerio de Ciencia e Innovación (MICINN), grants BFU2011-23978 and BFU2015-68149-R co-financed by the European Regional Development Fund].

ACKNOWLEDGMENTS

The authors thank Candido P. Segurado, for his contribution to the fly lines maintenance and the eye pictures collection, and Dr. Ismael Al-Ramahi and Dr. Maria de Haro, for loaning a set of pictures comprising different illumination conditions and genetic backgrounds, which were used for validation purposes. The authors also note that this work includes content that first appeared in the master's thesis of Diez-Hermano (2017).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2020.00516/full#supplementary-material>

- Dalal, N., and Triggs, B. (2005). "Histograms of oriented gradients for human detection," in *Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*, San Diego, CA.
- Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., and Fei-Fei, L. (2009). "Imagenet: A large-scale hierarchical image database," in *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL.
- Diez-Hermano, S. (2017). *Machine Learning Regularity Representation From Biological Patterns: A Case Study In A Drosophila Neurodegenerative Model*. Master thesis, Open University of Catalonia (UOC), Spain.
- Diez-Hermano, S., Valero, J., Rueda, C., Ganfornina, M. D., and Sanchez, D. (2015). An automated image analysis method to measure regularity in biological patterns: a case study in a *Drosophila* neurodegenerative model. *Mol. Neurodegener.* 10:9. doi: 10.1186/s13024-015-0005-z
- Fernandez-Funez, P., Nino-Rosales, M. L., de Gouyon, B., She, W. C., Luchak, J. M., Martinez, K., et al. (2000). Identification of genes that modify ataxin-1-induced neurodegeneration. *Nature* 408:6808. doi: 10.1038/35040584
- Ferri, C., Hernandez-Orallo, J., and Salido, M. A. (2003). "Volume under the ROC surface for multi-class problems. exact computation and evaluation of approximations," in *Proceedings of the 14th European Conference on Machine Learning*, València.
- Galasso, A., Cameron, C. S., Frenguelli, B. G., and Moffat, K. G. (2017). An AMPK-dependent regulatory pathway in tau-mediated toxicity. *Biol. Open* 6:10. doi: 10.1242/bio.022863
- Garcia-Lopez, A., Llamusi, B., Orzaez, M., Perez-Paya, E., and Artero, R. D. (2011). In vivo discovery of a peptide that prevents CUG-RNA hairpin formation and reverses RNA toxicity in myotonic dystrophy models. *Proc. Natl. Acad. Sci. U.S.A.* 108:29. doi: 10.1073/pnas.1018213108
- Giacinto, G., and Roli, F. (2001). Design of effective neural network ensembles for image classification purposes. *Image Vis. Comp.* 19:9. doi: 10.1016/S0262-8856(01)00045-2
- He, B. Z., Ludwig, M. Z., Dickerson, D. A., Barse, L., Arun, B., and Vilhjalmsson, B. J. (2014). Effect of genetic variation in a *Drosophila* model of diabetes-associated misfolded human proinsulin. *Genetics* 196:2. doi: 10.1534/genetics.113.157800
- Ioffe, S., and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv [Preprint]*, Available online at: <https://arxiv.org/abs/1502.03167> (accessed May 15, 2017).

- Iyer, J., Wang, Q., Le, T., Pizzo, L., Grönke, S., and Ambegaokar, S. (2016). Quantitative assessment of eye phenotypes for functional genetic studies using *Drosophila melanogaster*. *G3* 3:5. doi: 10.1534/g3.116.027060/-/DC1
- Jenny, A. (2011). Preparation of adult *Drosophila* eyes for thin sectioning and microscopic analysis. *J. Vis. Exp.* 54:2959. doi: 10.3791/2959
- Johnston, D. (2002). The art and design of genetic screens: *Drosophila melanogaster*. *Nat. Rev. Genet.* 3, 176–188.
- Jonshon, R. I., and Cagan, R. L. (2009). A Quantitative method to analyze *Drosophila* pupal eye patterning. *PLoS One* 4: e7008. doi: 10.1371/journal.pone.0007008.g001
- Karatzoglou, A., Smola, A., Hornik, K., and Zeileis, A. (2004). kernlab - An R Package for kernel methods in R. *J. Stat. Soft.* 11:9. doi: 10.18637/jss.v011.i09
- Kraus, O. Z., Ba, J. L., and Frey, B. J. (2016). Classifying and segmenting microscopy images with deep multiple instance learning. *Bioinformatics* 32:12. doi: 10.1093/bioinformatics/btw252
- Kuhn, M. (2008). Building predictive models in R using the caret package. *J. Stat. Soft.* 28:5. doi: 10.18637/jss.v028.i05
- Kuhn, M., Weston, S., Coulter, N., and Culp, M. (2015). *C50: C5.0 Decision Trees and Rule-Based Models R Package Version 0.1.0-24*. Available online at: <https://CRAN.R-project.org/package=C50> (accessed February 10, 2017).
- Lasagna-Reeves, C. A., de Haro, M., Hao, S., Park, J., Rousseaux, M. W., and Al-Ramahi, I. (2016). Reduction of Nuak1 decreases tau and reverses phenotypes in a tauopathy mouse model. *Neuron* 92:2. doi: 10.1016/j.neuron.2016.09.022
- LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521:7553. doi: 10.1038/nature14539
- Lenz, S., Karsten, P., Schulz, J. B., and Voigt, A. (2013). *Drosophila* as a screening tool to study human neurodegenerative diseases. *J. Neurochem.* 127:4. doi: 10.1111/jnc.12446
- Li, S., Xiabi, L., Ling, M., Chunwu, Z., Xinming, Z., and Yanfeng, Z. (2012). “Using HOG-LBP features and MMP learning to recognize imaging signs of lung lesions,” in *Proceedings of the 2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, Rome.
- Mishra, M., and Knust, E. (2013). Analysis of the *Drosophila* compound eye with light and electron microscopy. *Meth. Mol. Biol.* 935, 161–182. doi: 10.1007/978-1-62703-080-9_11
- Mouselimis, L. (2017). *OpenImageR: An Image Processing Toolkit. R Package Version 1.0.5*.
- Orrite, C., Gañán, A., and Rogez, G. (2009). “HOG based decision tree for facial expression classification,” in *Pattern Recognition and Image Analysis, 5524 (Lecture Notes in Computer Science)*, eds H. Araujo, A. Mendonça, A. Pinho, and M. Torres, (Berlin: Springer), 176–183.
- Pau, G., Fuchs, F., Sklyar, O., Boutros, M., and Huber, W. (2010). EBImage - an R package for image processing with applications to cellular phenotypes. *Bioinformatics* 26:7. doi: 10.1093/bioinformatics/btq046
- Po-Hsien, L., Shun-Feng, S., Ming-Chang, C., and Chih-Ching, H. (2015). “Deep Learning and its application to general image classification,” in *Proceedings of the 2015 International Conference on Informative and Cybernetics for Computational Social Systems (ICCSS)*, Chengdu. doi: 10.1186/s12859-017-1954-8
- R Core Team, (2018). *R: A Language And Environment For Statistical Computing*. Vienna: R Foundation for Statistical Computing.
- Reiter, L. T., Potocki, L., Chien, S., Gribskov, M., and Bier, E. (2001). A systematic analysis of human disease-associated gene sequences in *Drosophila melanogaster*. *Genome Res.* 11, 1114–1125. doi: 10.1101/gr.169101
- Roederer, K., Cozy, L., Anderson, J., and Kumar, J. P. (2005). Novel dominant-negative mutation within the six domain of the conserved eye specification gene *sine oculis* inhibits eye development in *Drosophila*. *Dev. Dyn.* 232:3. doi: 10.1002/dvdy.20316
- Rousseaux, M. W., de Haro, M., Lasagna-Reeves, C. A., De Maio, A., Park, J., and Vilanova-Velez, L. (2016). TRIM28 regulates the nuclear accumulation and toxicity of both alpha-synuclein and tau. *eLife* 5:19809. doi: 10.7554/eLife.19809
- Schroff, F., Criminisi, A., and Zisserman, A. (2008). “Object class segmentation using random forests,” in *Proceedings of the British Machine Vision Conference 2008*, Leeds.
- Sommer, C., and Gerlich, D. W. (2013). Machine learning in cell biology - teaching computers to recognize phenotypes. *J. Cell Sci.* 126:24. doi: 10.1242/jcs.123604
- Song, W., Smith, M. R., Syed, A., Lukacsovich, T., Barbaro, B. A., and Purcell, J. (2013). Morphometric analysis of Huntington’s disease neurodegeneration in *Drosophila*. *Meth. Mol. Biol.* 1017, 14–57. doi: 10.1007/978-1-62703-438-8_3
- Spanhol, F. A., Oliveira, L. S., Petitjean, C., and Heutte, L. (2016). “Breast cancer histopathological image classification using convolutional neural networks,” in *Proceedings of the 2016 International Joint Conference on Neural Networks (IJCNN)*, Vancouver, BC.
- Tarca, A., Carey, V. J., Chen, X. W., Romero, R., and Draghici, S. (2007). Machine learning and its applications to biology. *PLoS Comput. Biol.* 3:e116. doi: 10.1371/journal.pcbi.0030116
- Thaker, H. M., and Kankel, D. R. (1992). Mosaic analysis gives an estimate of the extent of genomic involvement in the development of the visual system in *Drosophila melanogaster*. *Genetics* 131, 883–894.
- Thomas, B. J., and Wassarman, D. A. (1999). A fly’s eye view of biology. *Trends. Genet.* 15, 184–190.
- Tianqi, C., Qiang, K., and Tong, H. (2015). *Mxnet: MXNet. R Package Version 0.9.4*. Available online at: <https://github.com/dmlc/mxnet/tree/master/R-package> (accessed March 3, 2017).
- Treisman, J. E. (2013). Retinal differentiation in *Drosophila*. *Interdiscip. Rev. Dev. Biol.* 2:4. doi: 10.1002/wdev.100
- Tyagi, V. (2019). A review on image classification techniques to classify neurological disorders of brain MRI. *IEEE ICICT* 1, 1–4. doi: 10.1109/ICICT46931.2019.8977658
- Vardi, Y., and Cun-Hui, Z. (2000). The multivariate L1-median and associated data depth. *PNAS* 97, 1423–1426. doi: 10.1073/pnas.97.4.1423
- Wangler, M. F., Yamamoto, S., and Bellen, H. J. (2015). Fruit flies in biomedical research. *Genetics* 199:3. doi: 10.1534/genetics.114.171785
- Winston, C., Joe, C., Allaire, J. J., Yihui, X., and McPherson, J. (2017). *Shiny: Web Application Framework for R. R Package Version 1.1.0*. Available online at: <https://CRAN.R-project.org/package=shiny> (accessed May 25, 2017).
- Yang, X., Yeo, S. Y., Hong, J. M., Wong, S. T., Tang, W. T., and Wu, Z. Z. (2016). A deep learning approach for tumor tissue image classification. *IASTED Biomed. Eng.* 832. doi: 10.2316/p.2016.832-025

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2020 Diez-Hermano, Ganfornina, Vegas-Lozano and Sanchez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.