Can *Drosophila melanogaster* tell who's who?

Jonathan Schneider[1,2,3], Nihal Murali[4], Graham W. Taylor[2,3,5], Joel D. Levine[1,3]*

1 University of Toronto, Mississauga, Ontario, Canada, **2** School of Engineering, University of Guelph, Guelph, Ontario, Canada, **3** Canadian Institute for Advanced Research, Toronto, Ontario, Canada, **4** Birla Institute of Technology and Science, Pilani, India, **5** Vector Institute for Artificial Intelligence, Toronto, Ontario, Canada

* joel.levine@utoronto.ca

Abstract

Drosophila melanogaster are known to live in a social but cryptic world of touch and odours, but the extent to which they can perceive and integrate static visual information is a hotly debated topic. Some researchers fixate on the limited resolution of *D. melanogaster*'s optics, others on their seemingly identical appearance; yet there is evidence of individual recognition and surprising visual learning in flies. Here, we apply machine learning and show that individual *D. melanogaster* are visually distinct. We also use the striking similarity of *Drosophila*'s visual system to current convolutional neural networks to theoretically investigate *D. melanogaster*'s capacity for visual understanding. We find that, despite their limited optical resolution, *D. melanogaster*'s neuronal architecture has the capability to extract and encode a rich feature set that allows flies to re-identify individual conspecifics with surprising accuracy. These experiments provide a proof of principle that *Drosophila* inhabit a much more complex visual world than previously appreciated.

Introduction

There is an increasing body of evidence that *Drosophila melanogaster* live in a surprisingly rich and complex world that includes group behaviour [1], communal learning [2], and recognition during aggressive behaviours [3]. This repertoire of social behaviour has often been assumed to be independent of visual recognition, as *Drosophila*'s compound eye was thought to have insufficient visual acuity to play a serious role. With approximately 850 lens units (ommatidia), each capturing a single point in space, *Drosophila*'s compound eye resolution is certainly low. Additionally, the level of detail, traditionally determined by the inter-ommatidial angle, renders anything but movement or regular patterns seemingly impossible to discern (Fig 1B).

However, recent physiological experiments [6] reveal that *D. melanogaster* can respond to details as fine as 1.16° as long as they are presented (to a tethered fly) at specific speeds. These speeds happen to coincide with *D. melanogaster*'s natural saccadic gait [7], which strongly suggests that naturally behaving *D. melanogaster* have much finer resolution than the inter-ommatidial angle of 4.8° [5] would imply. This hyper-acuity is found at the photoreceptor level (due to rhabdomere movement changing the angle of light reception), allowing most of the capacity of the visual network to be devoted to information processing. At this effective

hyper-acuity, and at socially relevant distances, the number of ommatidia and not the inter-ommatidial angle becomes the limiting factor (Fig 1). This acuity potentially puts them in the same visual league (albeit with lower resolution) as *Apis mellifera* [8], which has been able to, among other visual feats, identify individual human faces [9].

This spatio-temporal coding and increased visual acuity potentially explain recent studies which have shown that *D. melanogaster* can not only resolve other flies, but can also decode social meaning using vision (e.g. female choice of male phenotypes [10] and exposure to parasitoids [11]). Combined, these results open up the possibility that *Drosophila* uses vision to a much greater extent in object recognition, perhaps even using it to discriminate between species or sex (supplementing other olfactory cues that are known to convey this information [12]).

Even with *D. melanogaster*'s photoreceptor hyper-acuity [6], the image that is received is only around 29×29 units (or pixels; Fig 1). We wanted to know if there is enough absolute information contained in this low-resolution image to identify individuals from each other. One approach is to task Deep Convolutional Networks (DCNs) to differentiate individual *D. melanogaster*, as DCNs are engineered to learn, extract, and use any useful features found in the images. If there is enough individual-level variation for highly engineered DCNs, we would want to investigate the possibility that *D. melanogaster* also take this low resolution image and extract meaningful information out of it. Should individual flies prove visually unique and *D. melanogaster*'s visual network have enough capacity, vision could potentially play a role in identifying beyond species or sex, perhaps aiding in determining familiar or non-familiar conspecifics in social situations [3].

How a fly's visual system could extract meaning out of low resolution images is suggested by the highly structured and layered organization of *Drosophila*'s visual system (Fig 2C). At the input, the ommatidia are packed one by one, but their individually-tuned photoreceptors are arranged spatially to essentially convolve a 6-unit filter across the receptive field. The output of this photoreceptor filter is, in turn, the input for downstream medulla neurons that connect to several 'columns' of photoreceptor outputs. This filter-convolution, coupled with the use of the output from one filter as a 'feature map' for another layer is a hallmark of the engineered architectures of DCNs that dominate computer vision today (one such DCN is illustrated in Fig 2A). Just as DCNs can take low level image representations and encode
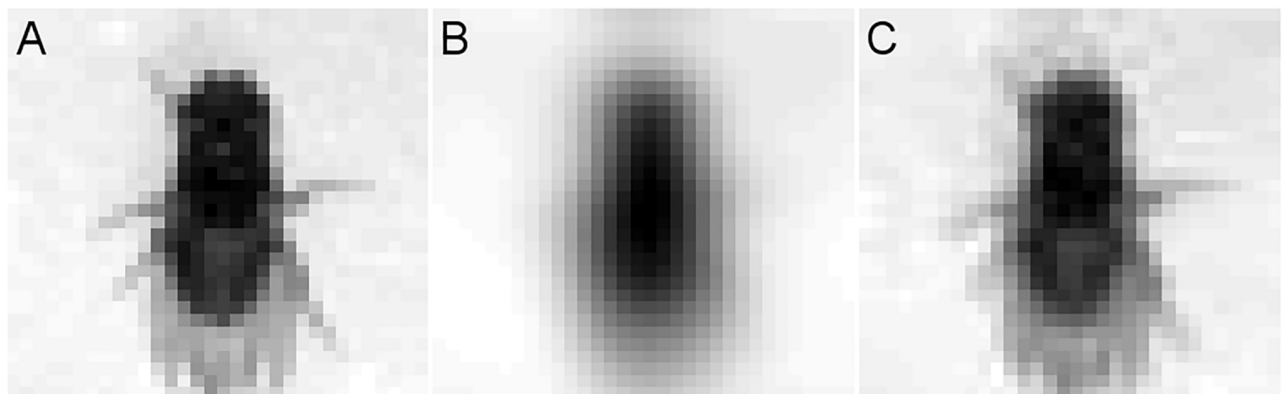
Fig 1. Theoretical visual acuity of *Drosophila melanogaster*. Image of *Drosophila melanogaster* represented after various theoretical bottlenecks. A: Image of a female *D. melanogaster* re-sized through a 32×32 bottleneck. B: The same image, but adjusted using AcuityView [4] for a viewing distance of 3 body lengths using the inter-ommatidial angle of 4.8˚ [5]. C: The same image and distance, but using a conservative estimate of the effective acuity determined by Juusola *et al.* [6] of approximately 1.5˚.

Fig 2. Our fly-eye merges engineered and biological architectures. Schematics of a 'standard' convolutional network, our fly-eye model, and a simplified visual connectome of *Drosophila*. A: Architecture of Zeiler and Fergus [13], receiving the original 181×181 pixel image of an individual *Drosophila melanogaster*. B: Our fly-eye model, receiving a 29×29 down-scaled image of an individual *Drosophila*, and showing connections betw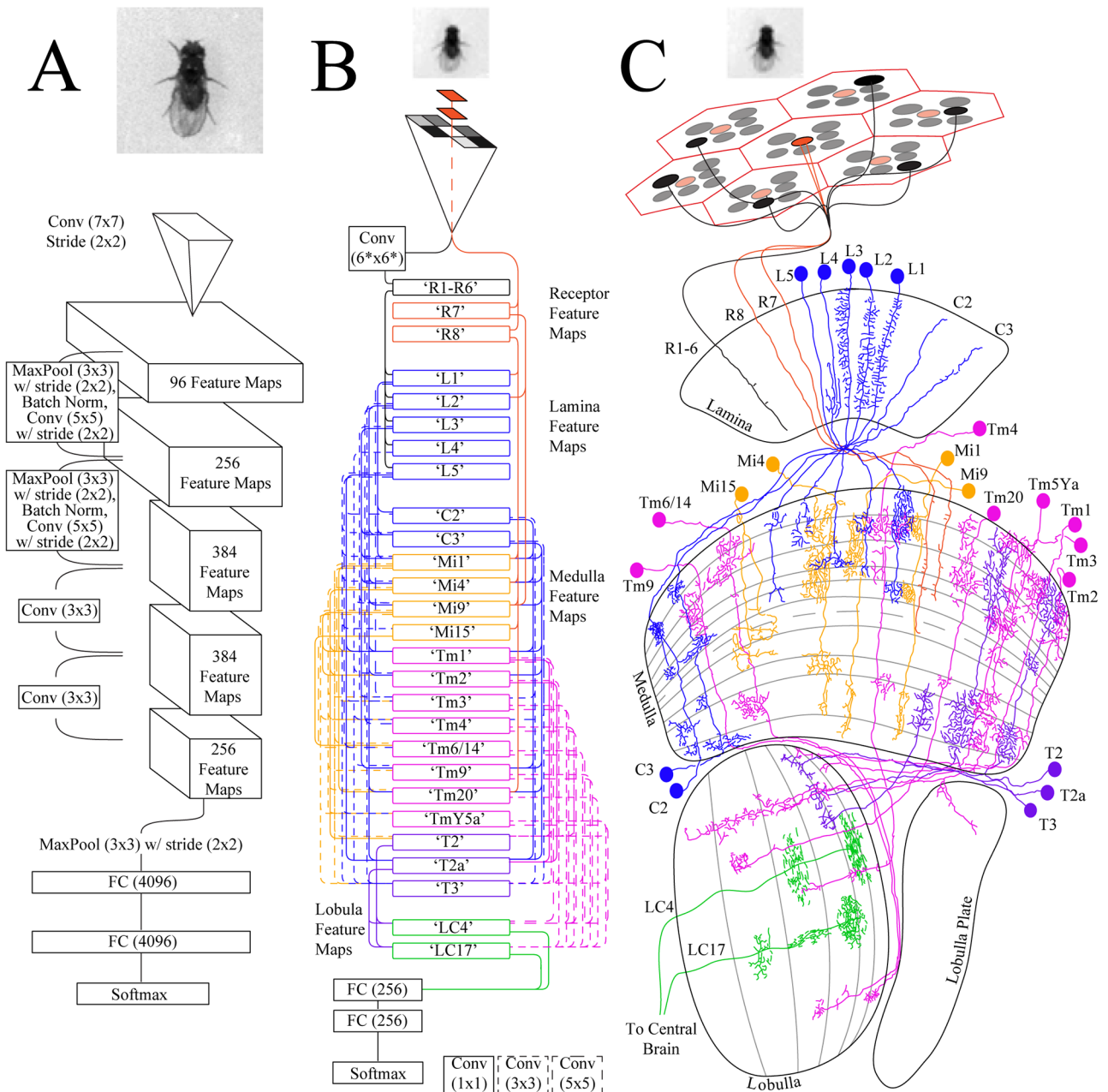een feature maps. The initial three feature maps are a custom 6-pixel convolutional filter ('R1-R6'; black pathway) and two 1×1 convolutional filters ('R7' and 'R8'; red pathway). All other convolutions are locally connected filters. See S1 Table for complete connectivity map. C: A simplified map of the fly visual circuit receiving the same down-scaled image of another *D. melanogaster*. The connections among the neurons implemented in our model are displayed, illustrating the connections and links within and between layers (adapted from [14] and [15]). See S2 Table for performance of these models on a traditional image-classification dataset.

them into a semantic representation, *D. melanogaster*'s visual system seems well-suited to similarly uncover semantic meaning in images.

In this work, we investigate whether *D. melanogaster* could theoretically categorize and recognize its complex visual environment. To determine how much absolute underlying visual

variation is available for social behaviour in *D. melanogaster*, we examine the ability of humans and human-inspired deep convolutional models to re-identify individual *D. melanogaster* across multiple days. To see whether *D. melanogaster* is *capable* of using this individual-level visual variation between flies, we investigate a model of *Drosophila*'s visual system in a conspecific re-identification paradigm. This study builds upon the behavioural results of conspecific information and physiological evidence of hyper-acute vision, and provides a proof of principle to dispel the oft-touted argument that *D. melanogaster*'s visual ability is limited to low-level object- and pattern-detection. Here, we present evidence that *D. melanogaster* may likely see and live in a much richer social environment than has been appreciated.

Materials and methods

Simplified *Drosophila* model eye

We implemented a virtual fly visual system using standard deep learning libraries (Keras). Our implementation uses approximately 25,000 artificial neurons, whereas *Drosophila* have approximately 60,000 neurons in each visual hemisphere [16]. We purposefully did not model neurons that are structurally suggestive to respond to movement, and therefore we were specifically limited to 'modular' neurons (with 1 neuron/column) throughout the medulla. The connections between neuronal types were extracted from published connectomes [17]. We imposed artificial hierarchy on our model eliminating self-connections between neuron 'subtypes' (i.e. no connections between L1 and L1, or L1 and L2), and while we allowed initial layers to feed into multiple downstream layers, we eliminated 'upstream' connections. The final lobula-like artificial neurons were modelled after Wu *et al.* [15], where the layers were ordered according to their axon penetration deeper into the system. Our ability to model *Drosophila*'s visual system is further limited to the connectivity, ignoring the sign (excitatory or inhibitory), as well as the neurons' intrinsic membrane properties. The ability to create more biologically realistic simulations will increase once these properties are discovered and integrated into the connectome. The model is illustrated in Fig 2B, beside the biological inspiration (Fig 2C). S1 Table depicts a complete connection map and hierarchy, and S2 Table shows comparative performance of this model on a traditional image-classification dataset. Additional details are provided in S1 Methods.

Fly data acquisition

D. melanogaster were housed at 25˚C on a 12-12 Light-Dark cycle. 10 males and 10 females were collected 1-4 hours after eclosion and housed separately. On the third day, post-eclosion flies were individually mouth pipetted into a circular acrylic arena (60mm diameter, 2mm high). These flies were illuminated with standard overhead LED bulbs and filmed in grayscale with a GRAS-20S4M for 15 minutes, with 16 frames/second. This was repeated for three consecutive days in total, which resulted in $14,400 \times 3$ images per fly. Each filming session was within 2 hours of ZT 8. Three independent datasets were acquired of 20 flies each.

Fly data processing

Each video was tracked using CTRAX [18], and the tracked position and orientation was used to localize the flies and orient the images so that the flies were always centred and facing 'up'. These images therefore contained all viewpoints of the flies during their behaviour in the arena: dorsal, ventral, and lateral. The training set for each biological replicate was constructed from days 1 and 2 equally, and consisted of the first 75% of the each fly's recording (12,240 frames). The validation set was the final 15% of the recording (2,160 frames). The test dataset

was the entire recording on day 3. Images were standardized by subtracting the mean and dividing by the standard deviation of the training set. For the ResNet18 [19] and Zeiler and Fergus [13] models, the original 181×181 image was either: (1) reduced to 33×33, centre-cropped to 29×29, and then re-sized to 224×224; or (2) re-sized to 256×256 and centre-cropped to 224×224 (effectively using the centre 158×158 pixels).

Human performance

A GUI program was written in MATLAB that presented a human observer with 3 viewpoints of an exemplar fly: dorsal, ventral, and sideways obtained from the first two days of filming. The observer was then asked to choose among 20 images (of the 20 flies) obtained from day 3, of which one belong to the exemplar (S3 and S4 Figs). Note that this is a compare/match setup rather than a learn/recall. The images were randomly re-sized through a 29×29 bottleneck.

Results

In this work, we wanted to see whether various architectures, both biologically rooted and not, could detect visual differences between flies across days (a notably non-human task). We acquired three rounds of images, each round having 10 males and 10 females, filmed for 3 consecutive days. Knowing that age/experience may slightly affect the morphology of flies, we experimented with training these networks on days 1 and 2, and testing their ability to re-identify the flies on day 3. We evaluated the efficient and high-capacity model (ResNet18 [19]), a model that rivals human representation (Zeiler and Fergus [13]), our fly-eye model, and human performance. These results are summarized in Table 1.

As a benchmark, we applied a ResNet18 architecture (see S1 Fig). This was the highest performing network architecture, achieving an F_1-score of 0.94 (with high resolution over the three datasets). While the average performance is good, we note certain problematic individual flies that become difficult to accurately re-identify across days (e.g. in biological replicate 2, fly 10's accuracy on day 3 was 37% with equal confusion between two other flies S4 Table). Forcing the images through a bottleneck (which ensures that the information content is similar to the reduced-resolution fly-eye model) decreased the F_1-score by 0.11 for ResNet18. The Zeiler

Table 1. Performance on *D. melanogaster* re-identification.

Model Name	Resolution[1] (pixels)	Accuracy (F_1 Score)[2]
ResNet18 [19]	158×158	0.9426 ± 0.0358
Zeiler and Fergus [13]	158×158	0.9373 ± 0.0365
Human Performance	158×158	0.1309
Zeiler and Fergus [13]	29×29	0.8549 ± 0.0778
ResNet18 [19]	29×29	0.8357 ± 0.0909
Our fly-eye	29×29	**0.7548 ± 0.1141**
Our fly-eye w/ random zoom[3]	29×29	**0.5486 ± 0.1316**
Human Performance	29×29	0.0829
Random Chance		0.05

Mean and standard deviation shown (n = 3 independent datasets)

[1] For Zeiler and Fergus and ResNet18, the "resolution" was a bottleneck (see Methods).

[2] An example of high precision / low recall can be seen in S4 Table as ID 10 is assigned to the right fly 94% of the time, but fly 10 is only correctly identified 37% of the time.

[3] The zoom was applied randomly without preserving aspect ratio (see S2 Fig).

https://doi.org/10.1371/journal.pone.0205043.t001

and Fergus [13] architecture was the most robust in terms of the bottleneck, only decreasing the F_1-score by 0.08, but did not achieve as high a performance as ResNet18.

The fly-eye model achieves a relatively high F_1 score of 0.75, which is not that much lower than the highly engineered ResNet18 (at low resolution). To eliminate the fly-eye model's ability to measure absolute size and shape and force relative feature extraction, we randomly resized the images (both training and testing) by up to 25% without preserving proportion (see S2 Fig for examples). Our virtual fly-eye outperformed humans, even without absolute size measures, achieving an F_1-score of 0.55 re-identifying conspecifics. We also note that the fly-eye model almost never mistakes a male for a female (F_1-score exceeding 0.99 when the re-identification IDs were collapsed by sex; S5–S7 Tables).

To establish a human performance baseline, we used volunteers to attempt the re-identification of flies (S3 and S4 Figs). This is an especially challenging task given the range of motion and viewpoints that an individual fly has in an unconstrained arena. As this task falls outside normal visual object recognition, we restricted volunteers to be very experienced "fly-pushing" scientists. Human performance was predictably low, but varied, with an average F_1-score of 0.11 (0.08 when the images were reduced to 29×29 pixels, and 0.13 when given the full resolution 181×181 images).

Discussion

Our results suggest that *Drosophila* have an innate capacity to extract semantic meaning from their visual surrounding, and even if we are still discovering *how Drosophila* could encode its world, we should not disregard its visual understanding.

The idea of recognizing broad semantic information decoded from a low resolution image is not new (for example, the success of CNNs on the 32×32 pixel CIFAR10 dataset and experiments in other downsampled datasets [20]). We also note that our fly-eye model performs relatively poorly on one such classification task (F_1 Score of 0.54 on CIFAR10; see S2 Table), which contains a wide range of scale and positional variance of the objects. One possible explanation for the inability to cope with variability in scale and size is that, unlike other architectures, *D. melanogaster*'s visual system maintains the input dimensions (the columnar medulla neurons). DCN's dimensionality reduction through 'tricks' like pooling layers and strided convolutions convey a larger position-invariance to low level feature detectors. Without them, our fly-eye model only does well when the distance to the object is fixed. It is tempting to posit that an individual can therefore have an innate and/or experience-dependent distance at which visual information is preferentially understood, and that this could be one of the determinants for social spacing and interaction distances [21, 22].

One prediction that arises from this model and its apparent ability to encode more than simple "looming" and "movement" visual cues is that the highest (furthest from input) level feature maps may correspond to semantically rich meaning in the visual system. These lobula neurons should then encode (among other information) complex object recognition categories and stimulating them should produce more than simple object-avoidance behaviours in *Drosophila*. While some lobula columnar neurons (like LC11) seem specialized for high-acuity small object motion detection [23], others appear to encode more complex information. These other LC neurons (like LC17), when stimulated, seem to provoke social-context dependent behaviours [15].

We are aware of other investigations using DCNs to classify insect species [24]. However, the most relevant study, which involves organism re-identification, only does so over short (<1 min) time frames (IDTracker2.0 [25]). Prior to this study, DCNs have only been effective on images that are temporally very close [25]. We observe a non-general loss of accuracy for specific flies, with some flies having an accuracy lower than 40% (S4 Table). This ability to re-

identify flies *across* days opens experimental possibilities, especially considering that this performance was achieved with static images (16fps yields around a thousand estimates of ID per minute, allowing high confidence in the parsimonious correct identification). This is in contrast to the human ability to re-identify flies, which at low resolutions is barely better than chance.

Clearly, all models can learn to re-identify flies to some extent, underscoring the individual-level variation in *D. melanogaster*. Re-identifying flies is in fact easier for DCNs than CIFAR10 (at least with centred images of flies acquired at the same distance). Even the model that rivals, in some sense, the representational performance of humans [26] does *ten times* better than humans. Why humans can't tell one fly from another is not clear. Regardless of whether it was evolutionarily beneficial to discriminate individual flies, humans do have incredible pattern detection abilities. It may simply be a lack of experience (although we attempted to address this by only using experienced *Drosophila* researchers as volunteers) or a more cryptic pattern-recognition 'blind-spot' of humans. In either case, these findings should spur new experiments to further understand the mechanisms of human vision and experience and how they fail in this case.

Machine learning practitioners are constantly pushing for deep networks to use more biologically-inspired design choices and training algorithms [27–29]. As they become more biologically realistic, neurobiologists could use these models to generate hypotheses in how information is processed in the visual system. We think *D. melanogaster* is well-suited to link these two fields to continue unravelling evolution's solution to visual processing. This new area offers a simple, genetically- and experimentally-tractable organism to peer into the workings of the visual system, which will no doubt uncover not just how *Drosophila*, but all of us, see the world.

Conclusion

These results help explain recent, traditionally controversial, findings that *Drosophila melanogaster* can resolve relatively detailed visual semantic meaning (e.g. female choice of males [2] and parasitoids exposure [11]). We show here that each *D. melanogaster* has visually distinguishable features that persist across days. This fact, combined with their hyper acuity [6] and theoretical capacity of their visual network, provides a proof of principle against the traditional belief that *Drosophila* only see blurry motion. In fact, *Drosophila* could have the ability to see and discriminate a surprisingly diverse visual world, in some cases, possibly better than we can.

Supporting information

S1 Methods. Additional details concerning the creation of the model (connectome), and the details/results of testing the model on a traditional image-classification dataset.
(PDF)

S1 Fig. ResNet18 architecture. Residual Networks with 18 'layers' [19] were constructed as depicted, using the improved block scheme proposed by [19] (top left inset).
(PDF)

S2 Fig. Examples of fly images. On the left are examples of the various viewpoints of the fly. The images indicated were run through the random non-proportion-preserving zoom to generate three examples of each.
(PDF)

S3 Fig. Human query: High resolution example. Participants were asked to re-identify the exemplar fly (top left), which exhibited three viewpoints: ventral, dorsal, and angled. They could request new images of said fly as many times as needed. The exemplar images were manually sorted into viewpoint categories and then randomly selected from Days 1 and 2. The other fly images (queries) were selected from Day 3.
(PNG)

S4 Fig. Human query: Low resolution example. Similar to the high-resolution experiment described above (S3 Fig), except all pictures were down sampled to 29×29.
(PNG)

S1 Table. Model fly-eye connectome. The table indicates the filter size for the locally-connected layers. No self connections were allowed (grey), or connections with 'higher' layers (blue). See S1 Methods for additional details.
(PDF)

S2 Table. Results of a simple vision task (CIFAR10). See S1 Methods for additional information about data processing and results.
(PDF)

S3 Table. Confusion matrix for ResNet18 biological replicate 1. Flies are ordered by sex (Purple = male, Yellow = Female), then by ascending size. Predictions are colour coded and weighted by percentage (correct predictions are indicated in orange, incorrect predictions are coloured cyan).
(PDF)

S4 Table. Confusion matrix for ResNet18 biological replicate 2. Note that fly 10 has a recall accuracy of 37%, and is almost equally confused for flies 2 and 7. Flies are ordered by sex (Purple = male, Yellow = Female), then by ascending size. Predictions are colour coded and weighted by percentage (correct predictions are indicated in orange, incorrect predictions are coloured cyan).
(PDF)

S5 Table. Confusion matrix for ResNet18 biological replicate 3. Flies are ordered by sex (Purple = male, Yellow = Female), then by ascending size. Predictions are colour coded and weighted by percentage (correct predictions are indicated in orange, incorrect predictions are coloured cyan).
(PDF)

S6 Table. Confusion matrix for fly-eye model biological replicate 1. Flies are ordered by sex (Purple = male, Yellow = Female), then by ascending size. Predictions are colour coded and weighted by percentage (correct predictions are indicated in orange, incorrect predictions are coloured cyan).
(PDF)

S7 Table. Confusion matrix for fly-eye model biological replicate 2. Flies are ordered by sex (Purple = male, Yellow = Female), then by ascending size. Predictions are colour coded and weighted by percentage (correct predictions are indicated in orange, incorrect predictions are coloured cyan).
(PDF)

S8 Table. Confusion matrix for fly-eye model biological replicate 3. Flies are ordered by sex (Purple = male, Yellow = Female), then by ascending size. Predictions are colour coded and

weighted by percentage (correct predictions are indicated in orange, incorrect predictions are coloured cyan).
(PDF)

Acknowledgments

Author Contributions

Conceptualization: Jonathan Schneider, Nihal Murali, Graham W. Taylor, Joel D. Levine.

Data curation: Jonathan Schneider, Graham W. Taylor.

Formal analysis: Jonathan Schneider, Graham W. Taylor.

Funding acquisition: Graham W. Taylor, Joel D. Levine.

Investigation: Jonathan Schneider.

Methodology: Jonathan Schneider, Nihal Murali, Graham W. Taylor.

Project administration: Jonathan Schneider, Graham W. Taylor.

Resources: Graham W. Taylor, Joel D. Levine.

Software: Jonathan Schneider, Nihal Murali.

Supervision: Jonathan Schneider, Graham W. Taylor, Joel D. Levine.

Validation: Jonathan Schneider, Nihal Murali.

Visualization: Jonathan Schneider.

Writing – original draft: Jonathan Schneider.

Writing – review & editing: Jonathan Schneider, Nihal Murali, Graham W. Taylor, Joel D. Levine.

References

1. Schneider J, Dickinson MH, Levine JD. Social structures depend on innate determinants and chemosensory processing in Drosophila. Proceedings of the National Academy of Sciences. 2012; 109(Supplement 2):17174–17179. https://doi.org/10.1073/pnas.1121252109

2. Battesti M, Pasquaretta C, Moreno C, Teseo S, Joly D, Klensch E, et al. Ecology of information: social transmission dynamics within groups of non-social insects. Proceedings of the Royal Society of London B: Biological Sciences. 2015; 282(1801):20142480. https://doi.org/10.1098/rspb.2014.2480

3. Yurkovic A, Wang O, Basu AC, Kravitz EA. Learning and memory associated with aggression in Drosophila melanogaster. Proceedings of the National Academy of Sciences. 2006; 103(46):17519–17524. https://doi.org/10.1073/pnas.0608211103

4. Caves EM, Johnsen S. AcuityView: An r package for portraying the effects of visual acuity on scenes observed by an animal. Methods in Ecology and Evolution. 2018; 9(3):793–797. https://doi.org/10.1111/2041-210X.12911

5. Land MF. Visual acuity in insects. Annual review of entomology. 1997; 42(1):147–177. https://doi.org/10.1146/annurev.ento.42.1.147 PMID: 15012311

6. Juusola M, Dau A, Song Z, Solanki N, Rien D, Jaciuch D, et al. Microsaccadic sampling of moving image information provides Drosophila hyperacute vision. eLife. 2017; 6. https://doi.org/10.7554/eLife.26117 PMID: 28870284

7. Geurten BR, Jähde P, Corthals K, Göpfert MC. Saccadic body turns in walking Drosophila. Frontiers in behavioral neuroscience. 2014; 8:365. https://doi.org/10.3389/fnbeh.2014.00365 PMID: 25386124

8. Caves EM, Brandley NC, Johnsen S. Visual Acuity and the Evolution of Signals. Trends in ecology & evolution. 2018;. https://doi.org/10.1016/j.tree.2018.03.001

9. Dyer AG, Neumeyer C, Chittka L. Honeybee (Apis mellifera) vision can discriminate between and recognise images of human faces. Journal of Experimental Biology. 2005; 208(24):4709–4714. https://doi.org/10.1242/jeb.01929 PMID: 16326952

10. Mery F, Varela SA, Danchin É, Blanchet S, Parejo D, Coolen I, et al. Public versus personal information for mate copying in an invertebrate. Current Biology. 2009; 19(9):730–734. https://doi.org/10.1016/j.cub.2009.02.064 PMID: 19361993

11. Kacsoh BZ, Bozler J, Bosco G. Drosophila species learn dialects through communal living. bioRxiv. 2017; p. 206920.

12. Billeter JC, Atallah J, Krupp JJ, Millar JG, Levine JD. Specialized cells tag sexual and species identity in Drosophila melanogaster. Nature. 2009; 461(7266):987. https://doi.org/10.1038/nature08495 PMID: 19829381

13. Zeiler MD, Fergus R. Visualizing and understanding convolutional networks. In: European conference on computer vision. Springer; 2014. p. 818–833.

14. Fischbach KF, Dittrich A. The optic lobe of Drosophila melanogaster. I. A Golgi analysis of wild-type structure. Cell and tissue research. 1989; 258(3):441–475. https://doi.org/10.1007/BF00218858

15. Wu M, Nern A, Williamson WR, Morimoto MM, Reiser MB, Card GM, et al. Visual projection neurons in the Drosophila lobula link feature detection to distinct behavioral programs. Elife. 2016; 5. https://doi.org/10.7554/eLife.21022

16. Raghu SV, Claussen J, Borst A. Neurons with GABAergic phenotype in the visual system of Drosophila. Journal of Comparative Neurology. 2013; 521(1):252–265. https://doi.org/10.1002/cne.23208 PMID: 22886821

17. Takemura Sy, Nern A, Chklovskii DB, Scheffer LK, Rubin GM, Meinertzhagen IA. The comprehensive connectome of a neural substrate for 'ON'motion detection in Drosophila. Elife. 2017; 6.

18. Branson K, Robie AA, Bender J, Perona P, Dickinson MH. High-throughput ethomics in large groups of Drosophila. Nature Methods. 2009; 6(6):451–457. https://doi.org/10.1038/nmeth.1328 PMID: 19412169

19. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition; 2016. p. 770–778.

20. Chrabaszcz P, Loshchilov I, Hutter F. A downsampled variant of ImageNet as an alternative to the CIFAR datasets. arXiv preprint arXiv:170708819. 2017;.

21. Schneider J, Levine J. Automated identification of social interaction criteria in Drosophila melanogaster. Biology letters. 2014; 10(10):20140749. https://doi.org/10.1098/rsbl.2014.0749 PMID: 25354920

22. Simon AF, Chou MT, Salazar ED, Nicholson T, Saini N, Metchev S, et al. A simple assay to study social behavior in Drosophila: measurement of social space within a group. Genes, Brain and Behavior. 2012; 11(2):243–252. https://doi.org/10.1111/j.1601-183X.2011.00740.x

23. Keleş MF, Frye MA. Object-detecting neurons in Drosophila. Current Biology. 2017; 27(5):680–687. https://doi.org/10.1016/j.cub.2017.01.012 PMID: 28190726

24. Ding W, Taylor G. Automatic moth detection from trap images for pest management. Computers and Electronics in Agriculture. 2016; 123:17–28. https://doi.org/10.1016/j.compag.2016.02.003

25. Romero-Ferrero F, Bergomi MG, Hinz R, Heras FJ, deÂ Polavieja GG. idtracker. ai: Tracking all individuals in large collectives of unmarked animals. arXiv preprint arXiv:180304351. 2018;.

26. Cadieu CF, Hong H, Yamins DL, Pinto N, Ardila D, Solomon EA, et al. Deep neural networks rival the representation of primate IT cortex for core visual object recognition. PLoS computational biology. 2014; 10(12):e1003963. https://doi.org/10.1371/journal.pcbi.1003963 PMID: 25521294

27. Bengio Y, Lee DH, Bornschein J, Mesnard T, Lin Z. Towards biologically plausible deep learning. arXiv preprint arXiv:150204156. 2015;.

28. Sabour S, Frosst N, Hinton GE. Dynamic routing between capsules. In: Advances in Neural Information Processing Systems; 2017. p. 3859–3869.

29. Tripp B. A deeper understanding of the brain. NeuroImage. 2017;. https://doi.org/10.1016/j.neuroimage.2017.12.079 PMID: 29292137