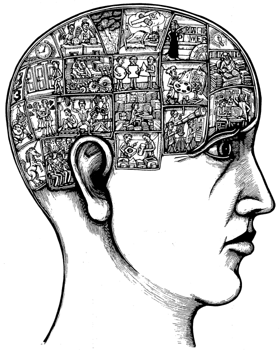
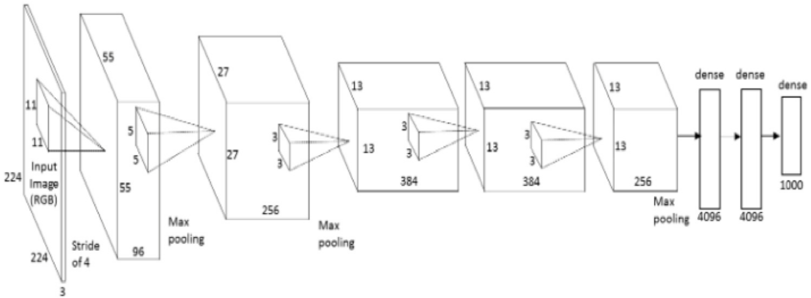
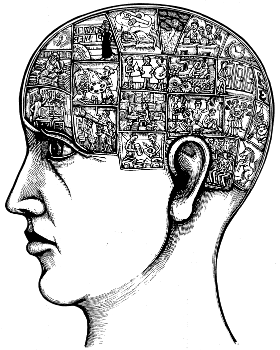
39th Annual Meeting of the Cognitive Science Society





# Functionally localized representations contain distributed information: insight from simulations of deep convolutional neural networks


**Nicholas Blauch (nblauch@umass.edu)**
University of Massachusetts, Amherst
Center for the Neural Basis of Cognition, Pittsburgh PA

**Elissa Aminoff**
Department of Psychology, Fordham University
Center for the Neural Basis of Cognition, Pittsburgh PA

**Michael J. Tarr**
Department of Psychology, Carnegie Mellon University
Center for the Neural Basis of Cognition, Pittsburgh PA

# Acknowledgments

CNBC
CENTER FOR THE NEURAL BASIS OF COGNITION

Michael J. Tarr
Elissa Aminoff
Ying Yang

UMASS AMHERST

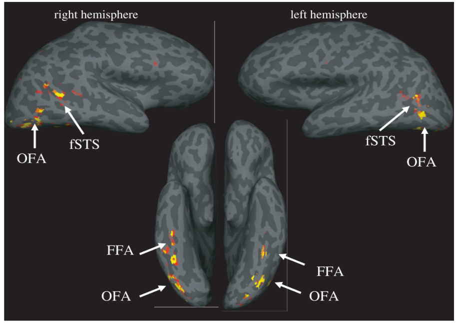
Rosie A. Cowell
David E. Huber

## Big question for this talk

Is the neural basis of face recognition distinct from that of general visual recognition?

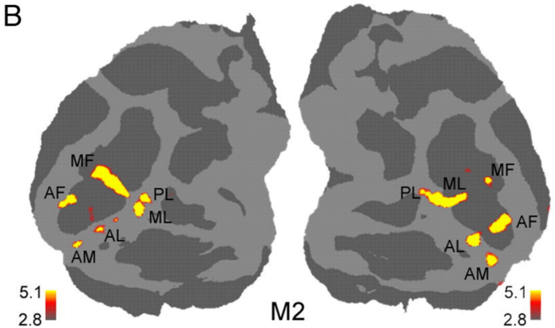
# Functional localization of face processing

Human cortical regions selective to faces (faces > objects, p<.00001)



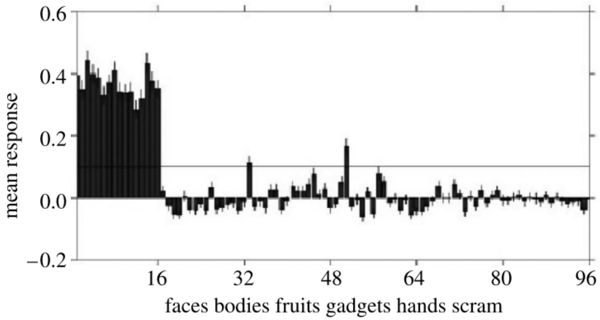

Kanwisher & Yovel, 2006

Macaque monkey "face patches" selective to faces vs. objects




Tsao, Moeller, Freiwald 2008

Macaque face patches show enhanced mean neural response to faces




Tsao,, Freiwald, Tootell & Livingstone, 2006

Face selectivity exists in stable locations in macaques as early as 200 days and overlaps with foveal and curvature selectivity present in the first postnatal weeks (Livingstone et. al, 2017; Arcaro & Livingstone, 2017)

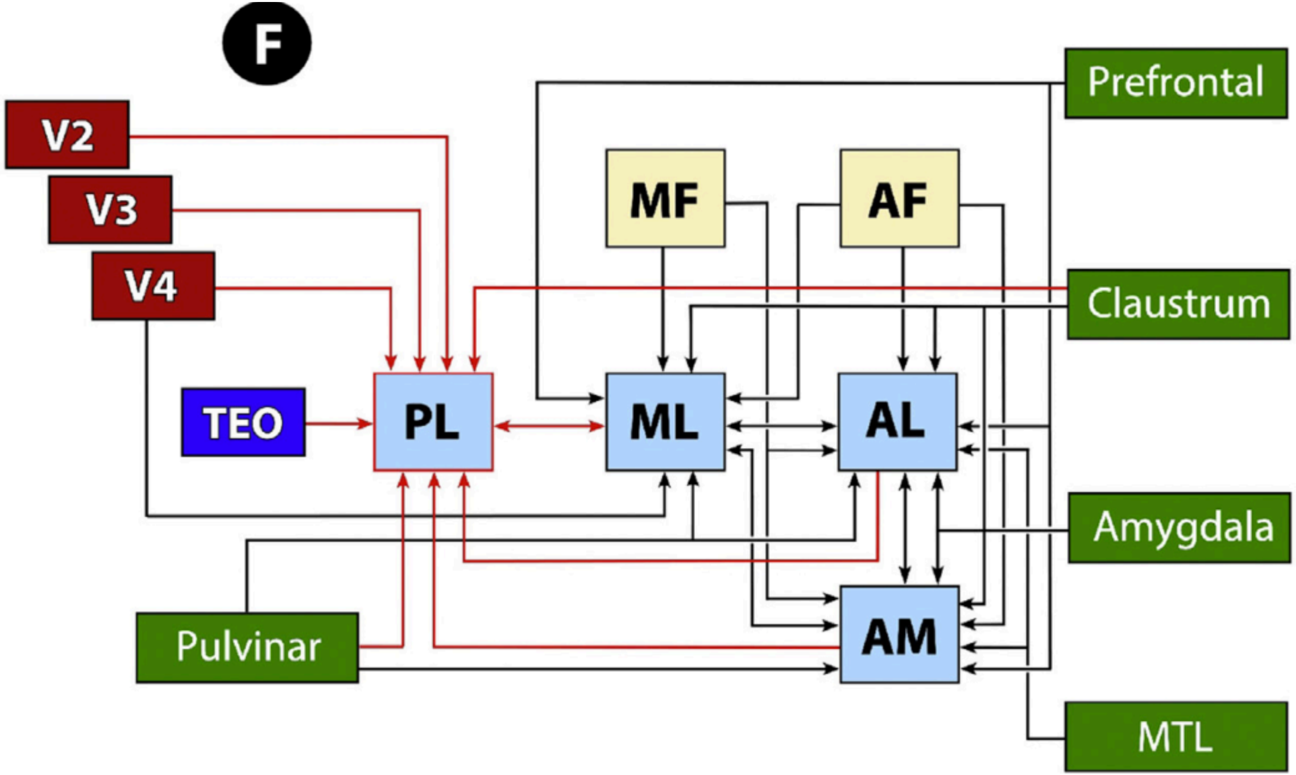
Face selectivity seems to be driven by low-level features and refined by perceptual experience recognizing faces

Tracer injections in face patches reveal greater connectivity with each other than than other nearby regions (Grimaldi, Saleem, Tsao, 2016)
-> functional sub-hierarchy honed by evolution?

CNBC CENTER FOR THE NEURAL BASIS OF COGNITION

UMASS AMHERST

# Proposed wiring diagram of a functional sub-hierarchy for face processing




Grimaldi, Saleem, Tsao (2016)

# Are face-selective regions just part of a distributed code?

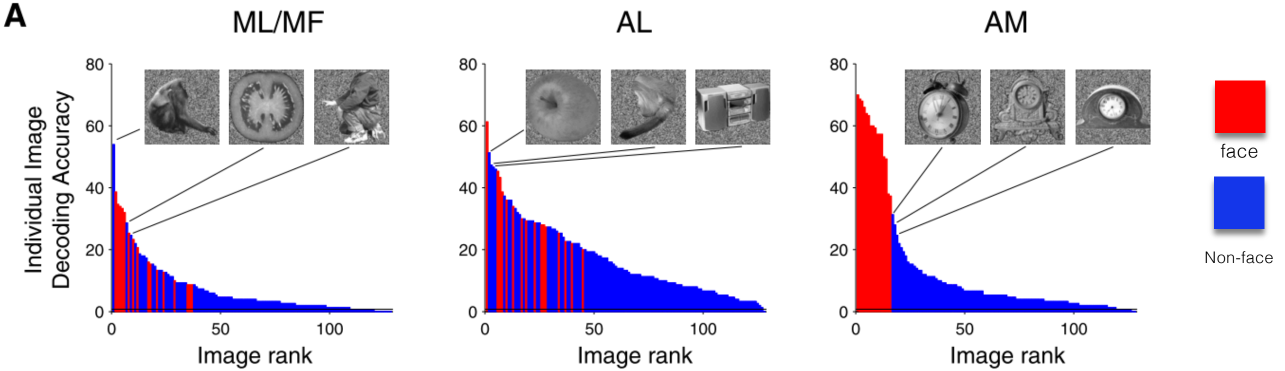
Multivariate pattern analysis (MVPA) of human face-selective areas supports non-face discrimination, and MVPA can discriminate faces without face-selective areas (Haxby et. al, 2001)

MVPA of all macaque “face patches” supports discrimination of most non-face images (Meyers et. al, 2015)

However, there are clear advantages for faces, especially in the most anterior patch AM

Seems unlikely that processing in face selective areas is a full-fledged module gated by earlier face-detection (e.g. Tsao & Livingstone, 2008)

Preferences in mean firing rate and information suggest a form of un-gated “specialization” or marginal “modularity”




Meyers et. al, 2015

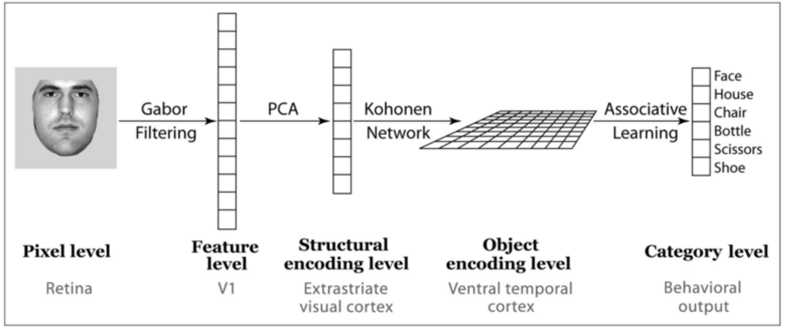
CNBC
CENTER FOR THE NEURAL BASIS OF COGNITION

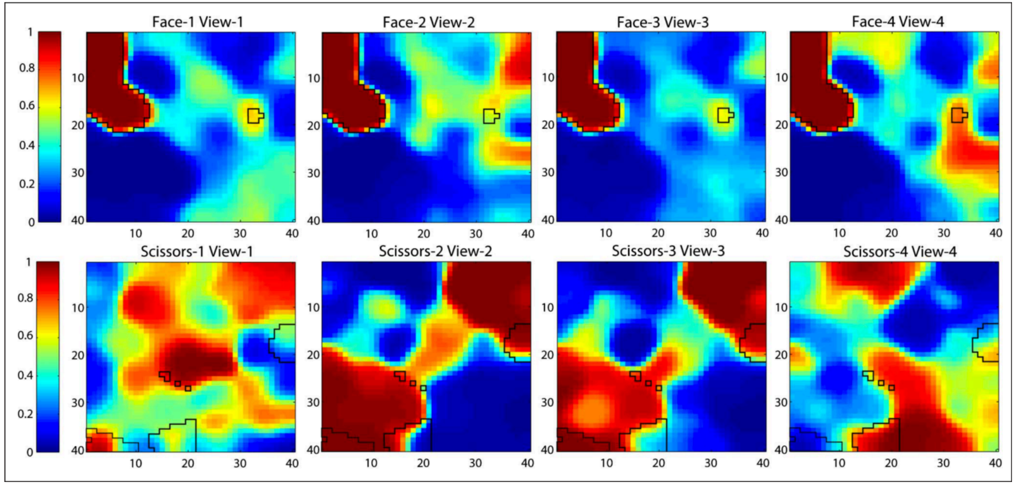
UMASS
AMHERST

# A model of distributed representation produces face-selectivity

Visual representations stored in a **domain-general** self-organizing Kohonen grid naturally develop strong view-independent face preferences compared to other categories (Cowell & Cottrell, 2013)

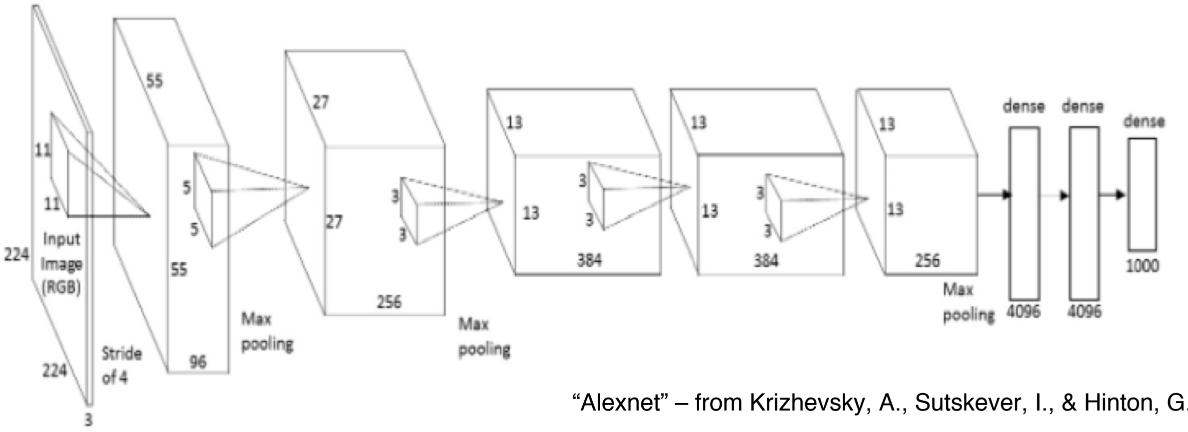
- However this model cannot perform face individuation – all individuals look basically the same!
- Not constraining to individuation belies the need for specialized circuitry for faces above other object categories for which individuation is less important




Cowell & Cottrell, 2013

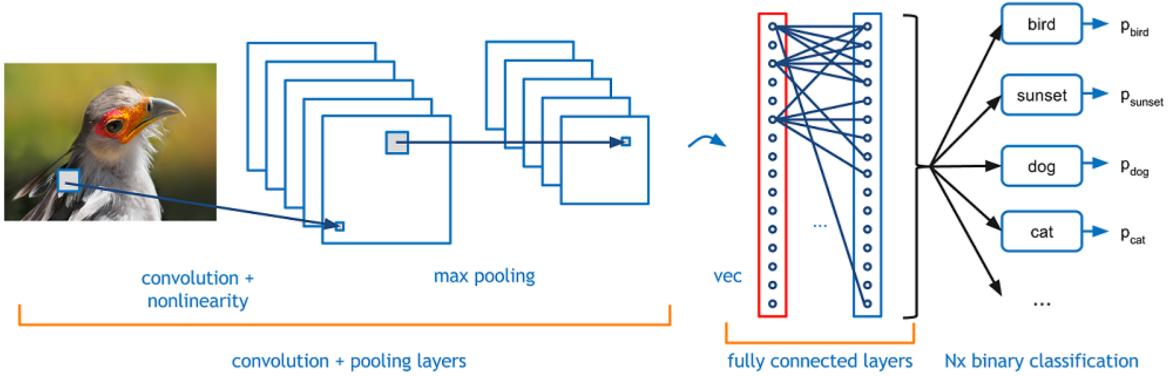
# Deep convolutional neural networks can perform human-level object recognition and face individuation




"Alexnet" – from Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012).

- Do functionally localized DCNN representations yield category-distributed information?

# Deep convolutional neural networks for visual categorization

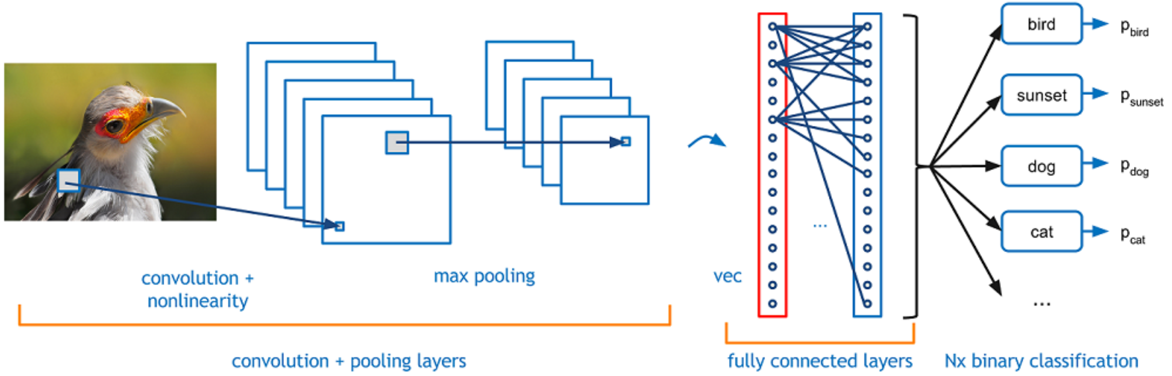



- DCNNs are the state-of-the-art in computer vision applications relating to high-level visual knowledge, such as categorization
- They consist of a hierarchy of processing layers and learn features through training on millions of images
- DCNN representations provide the best fits of neural data related to object recognition (Yamins et. al, 2014; Güclü & van Gervan, 2015)
- Impressive behavioral achievements, biological-inspiration, and ability to predict neural responses make them a powerful class of models for cognitive science



Graphic provided by Adit Deshpande's guide to understanding CNNs
https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

# Deep convolutional neural networks for visual categorization

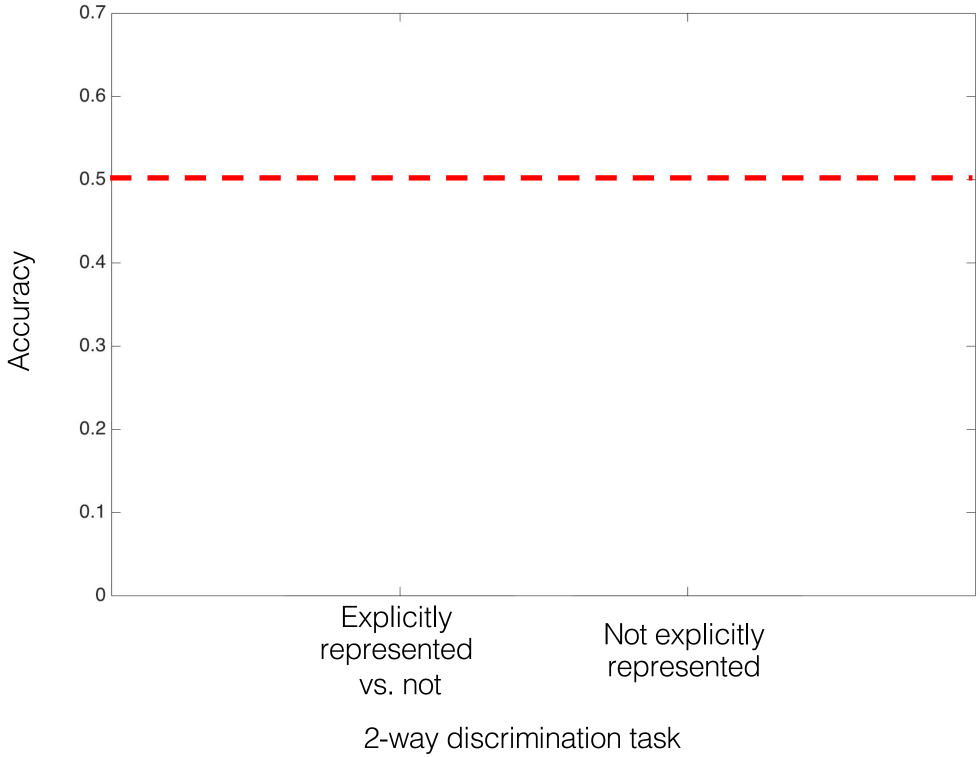



How much information do **localized categorical representations** contain about other categories?

Graphic provided by Adit Deshpande's guide to understanding CNNs
https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

# **Results**: decoding single-unit localized categorical representations in Alexnet

Using 10-fold (80% training) cross-validated SVM classification with
bootstrapped pairs across a sample of 20 ImageNet categories



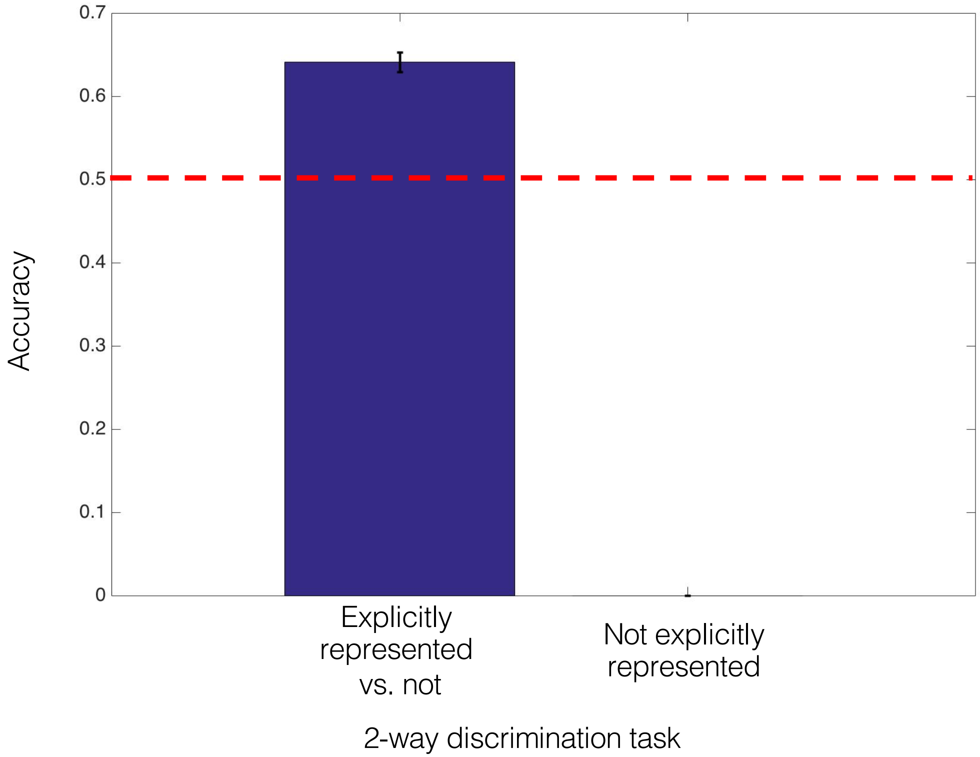

Blauch, Aminoff, Tarr (2017)

CNBC CENTER FOR THE NEURAL BASIS OF COGNITION

UMASS AMHERST

# **Results**: decoding single-unit localized categorical representations in Alexnet

Using 10-fold (80% training) cross-validated SVM classification with bootstrapped pairs across a sample of 20 ImageNet categories



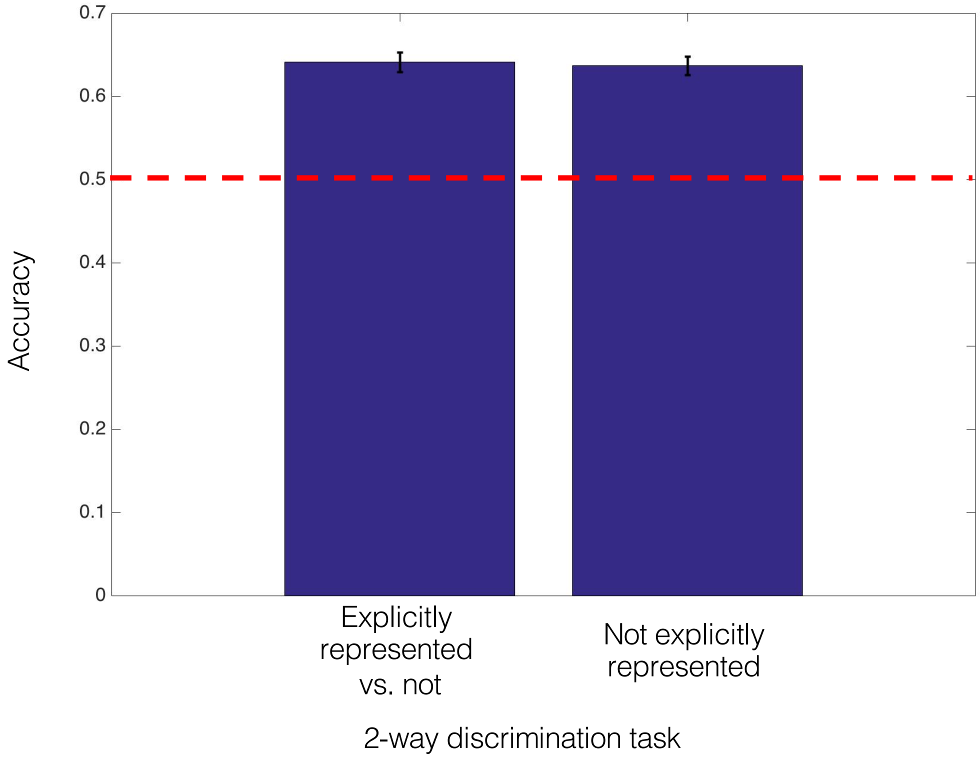

Blauch, Aminoff, Tarr (2017)

CNBC
CENTER FOR THE NEURAL BASIS OF COGNITION

UMASS AMHERST

# **Results**: decoding single-unit localized categorical representations in Alexnet

Using 10-fold (80% training) cross-validated SVM classification with bootstrapped pairs across a sample of 20 ImageNet categories

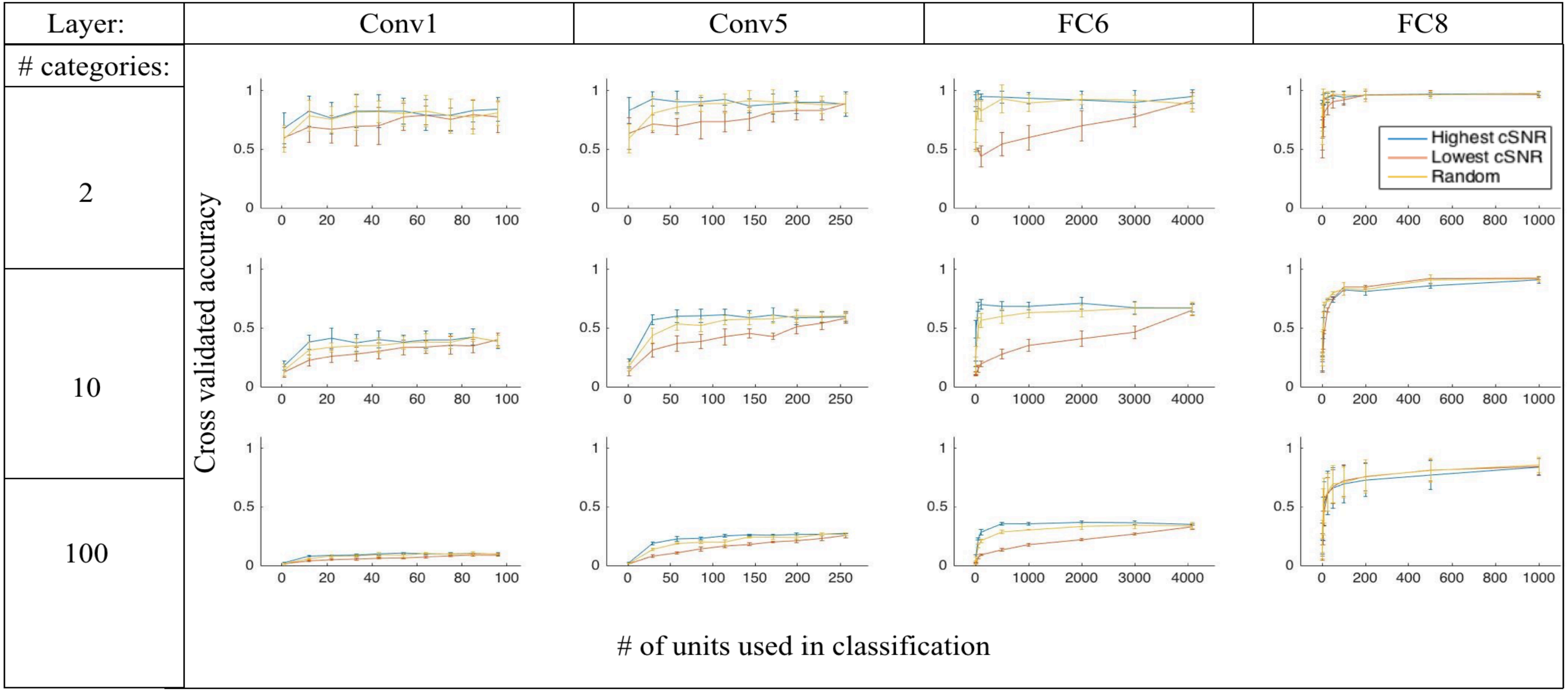



Blauch, Aminoff, Tarr (2017)

CNBC CENTER FOR THE NEURAL BASIS OF COGNITION

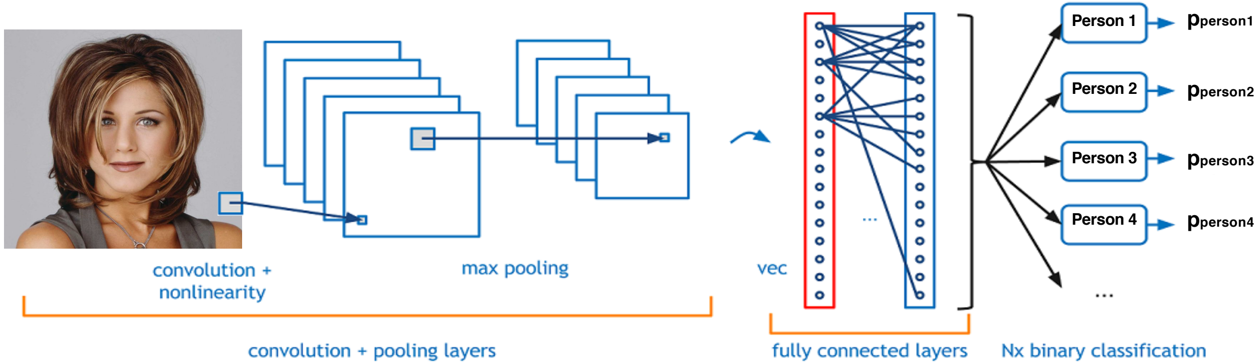
UMASS AMHERST

# Decoding populations of units across layers in Alexnet

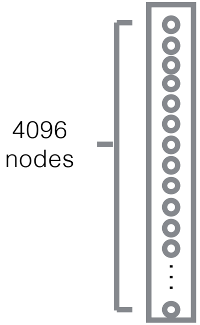



Blauch, Aminoff, Tarr (2017)

# DCNNs for face individuation




**A computational face "module"**



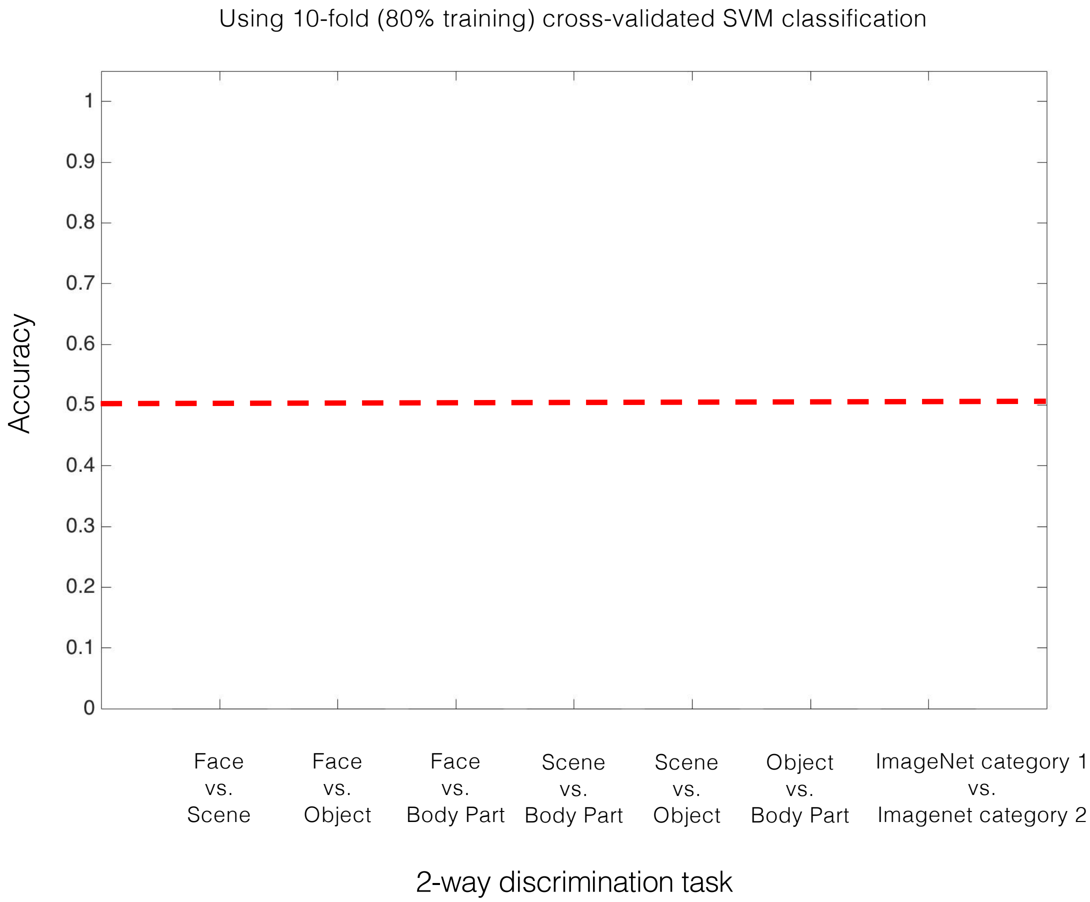

- **VGG-face** is a state-of-the-art DCNN for face individuation, trained to individuate 2622 faces at >90% accuracy (Parki, Vedaldi, and Zimmerman, 2015)
- The last fully-connected layer contains 4096 high-level features trained specifically for face individuation
    - We use these representations as a model of high-level face-specialized representations, e.g. a "modular" hypothesis about FFA or face patch AL

CNBC CENTER FOR THE NEURAL BASIS OF COGNITION

DCNN architecture cartoon adapted from Adit Deshpande's guide to understanding CNNs
https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks/

UMASS AMHERST

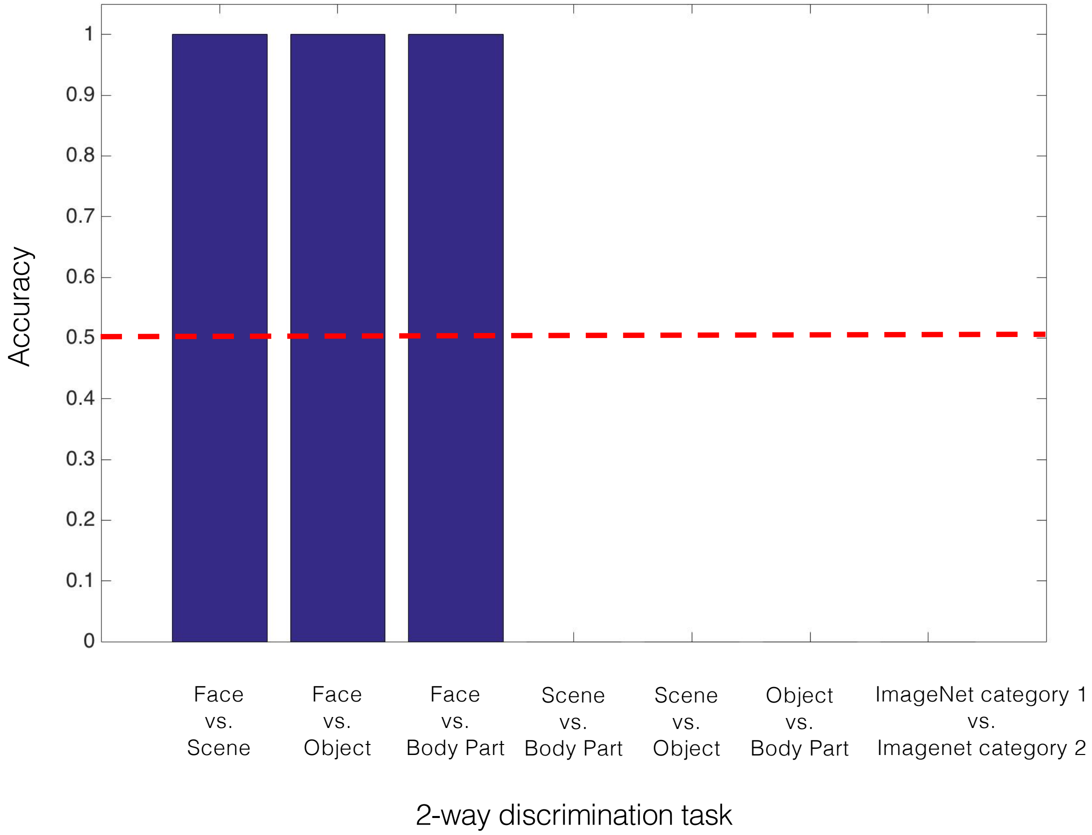
# **Results**: decoding a computational face module




Blauch, Aminoff, Tarr (2017)

# **Results**: decoding a computational face module

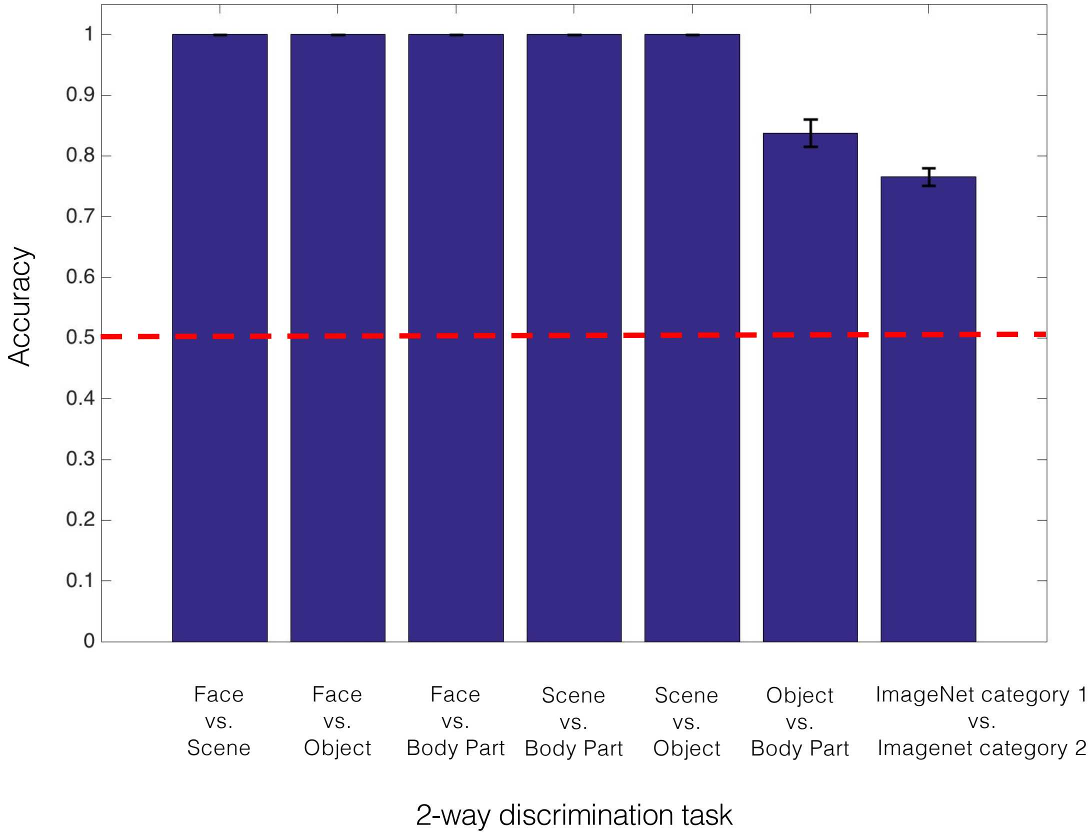
Using 10-fold (80% training) cross-validated SVM classification




Blauch, Aminoff, Tarr (2017)

# **Results**: decoding a computational face module

Using 10-fold (80% training) cross-validated SVM classification




Blauch, Aminoff, Tarr (2017)

CNBC CENTER FOR THE NEURAL BASIS OF COGNITION

UMASS AMHERST

# Conclusions

- The brain is highly selective for visual faces (e.g. Kanwisher & Yovel, 2006)
- Face-selective areas seem to still process non-face inputs (Haxby et. al, 2001; Meyers et. al, 2015)
- A **distributed topographic representation** produces the characteristic evidence of functionally localized activation (Cowell & Cottrell, 2013) but is shallow and cannot perform face individuation
- We show that deep **functionally localized representations** sufficient for human-level object categorization and face individuation produce categorically distributed information
  - Information can be deceiving! Information != function. Must build models to get at function.
- Current neuroscience evidence is thus consistent with both distributed and localized processing of faces
- Due to behavioral requirements and stimulus properties, we suspect that the true face processing scheme is moderately modular – containing patches highly specialized/trained for faces but not formally gated and thus adaptable to visual experience

# Discussion: the origins of functional localization of face processing

- Self-organizing principles lead to face-selectivity based on within-category feature similarity alone (Cowell & Cottrell, 2013)
- Innate face looking preference (e.g. Johnson et. al, 1991; Reid et. al, 2017) creates input bias, causing greater tuning to faces. This preference may be mediated by innate subcortical circuitry (Johnson, 2005).
- Regularization over connection length leads to greater modularization (Clune, Mouret, Lipson, 2013)
- Expertise with individuating faces causes top-down refinement of face-selective circuitry (e.g. Tarr & Gauthier, 2000)
  - Could be the basis of face specialization, or face-specialization could be the basis of other categories of expertise eliciting high activation in FFA (e.g. birds, greebles, cars)
    - Finer-grained representations developed for face recognition may be usefully coopted for other face-like categories requiring expertise
- Next steps – build models to test the behavioral and "neuroscientific" effects of these various factors

# References


- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). ImageNet Classification with Deep Convolutional Neural Networks. Advances In Neural Information Processing Systems, 1–9. https://doi.org/http://dx.doi.org/10.1016/j.protcy.2014.09.007
- Spiridon, M., & Kanwisher, N. (2002). How distributed is visual category information in human occipito-temporal cortex? An fMRI study. Neuron, 35(6), 1157–1165. https://doi.org/10.1016/S0896-6273(02)00877-2
- Haxby, J. V, Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science, 293(September), 2425–2430. https://doi.org/10.1126/science.1063736
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. The Journal of Neuroscience : The Official Journal of the Society for Neuroscience, 17(11), 4302–11. https://doi.org/10.1098/Rstb.2006.1934
- Vedaldi, A., & Lenc, K. (2015). MatConvNet. Proceedings of the 23rd ACM International Conference on Multimedia - MM '15, 689–692. https://doi.org/10.1145/2733373.2807412
- Nguyen, M. N., Nishimaru, H., Matsumoto, J., Le, Q. Van, & Hori, E. (2016). Population Coding of Facial Information in the Monkey Superior Colliculus and Pulvinar, 10(December). https://doi.org/10.3389/fnins.2016.00583
- Grimaldi, P., Saleem, K. S., Tsao, D., Grimaldi, P., Saleem, K. S., & Tsao, D. (2016). Anatomical Connections of the Functionally Defined "' Face Patches '" in the Macaque Monkey Article Anatomical Connections of the Functionally Defined "' Face Patches '" in the Macaque Monkey. Neuron, 90(6), 1325–1342. https://doi.org/10.1016/j.neuron.2016.05.009
- Chang, L., & Tsao, D. Y. (2017). The Code for Facial Identity in the Primate Brain. Cell, 169(6), 1013–1020.e14. https://doi.org/10.1016/j.cell.2017.05.011
- Meyers, E. M., Borzello, M., Freiwald, W. A., & Tsao, D. (2015). Intelligent Information Loss: The Coding of Facial Identity, Head Pose, and Non-Face Information in the Macaque Face Patch System. Journal of Neuroscience, 35(18), 7069–7081. https://doi.org/10.1523/JNEUROSCI.3086-14.2015
- Hanson, S. J., & Schmidt, A. (2011). High-resolution imaging of the fusiform face area (FFA) using multivariate non-linear classifiers shows diagnosticity for non-face categories. NeuroImage, 54(2), 1715–1734. https://doi.org/10.1016/j.neuroimage.2010.08.028
- Kanwisher, N. (2010). Functional specificity in the human brain: a window into the functional architecture of the mind. Proceedings of the National Academy of Sciences of the United States of America, 107(25), 11163–11170. https://doi.org/10.1073/pnas.1005062107
- Tsao, D. Y., & Livingstone, M. S. (2008). Mechanisms of Face Perception. Annual Reviews Neuroscience, 31, 411–37. https://doi.org/10.1146/annurev.neuro.30.051606.094238


- Chang, L., & Tsao, D. Y. (2017). The Code for Facial Identity in the Primate Brain. Cell, 169(6), 1013–1020.e14. https://doi.org/10.1016/j.cell.2017.05.011


- Grimaldi, P., Saleem, K. S., & Tsao, D. (2016). Anatomical Connections of the Functionally Defined "' Face Patches '" in the Macaque Monkey. *Neuron*, *90*(6), 1325–1342. https://doi.org/10.1016/j.neuron.2016.05.009
- Meyers, E. M., Borzello, M., Freiwald, W. A., & Tsao, D. (2015). Intelligent Information Loss: The Coding of Facial Identity, Head Pose, and Non-Face Information in the Macaque Face Patch System. *Journal of Neuroscience*, *35*(18), 7069–7081. https://doi.org/10.1523/JNEUROSCI.3086-14.2015
- Gauthier, I. (2017). The Quest for the FFA led to the Expertise Account of its Specialization. *Arxiv.org*
- Johnson, M. H., Dziurawiec, S., Ellis, H., & Morton, J. (1991). Newborns' preferential tracking of face-like stimuli and its subsequent decline. *Cognition*, *40*(1–2), 1–19. https://doi.org/10.1016/0010-0277(91)90045-6
- Reid, V. M., Dunn, K., Young, R. J., Amu, J., Donovan, T., Reissland, N. (2017). The Human Fetus Preferentially Engages with Face-like Visual Stimuli. *Current Biology*, *27*(12), 1825–1828.e3. https://doi.org/10.1016/j.cub.2017.05.044
- Adit Deshpande's GitHub: A Beginner's Guide To Understanding Convolutional Neural Networks https://adeshpande3.github.io/adeshpande3.github.io/A-Beginner%27s-Guide-To-Understanding-Convolutional-Neural-Networks

Thanks!

Questions?

CNBC
CENTER FOR THE NEURAL BASIS OF COGNITION

UMASS
AMHERST