Matt Welch are supported by NIH/NIAID grant R01 AI109044

WEB RESOURCES

U.S. Food & Drug Administration, https://www.fda.gov/

REFERENCES

Allan, E.J., Hoischen, C., and Gumpert, J. (2009). Bacterial L-forms. Adv. Appl. Microbiol. 68, 1–39.

Brown, G.W., King, G., and Sugiyama, H. (1970). Penicillin-Lysozyme Conversion of Clostridium botulinum Types A and E into Protoplasts and Their Stabilization as L-Form Cultures. J. Bacteriol. 104, 1325–1331

Domingue, G.J., Sr., and Woody, H.B. (1997). Bacterial persistence and expression of disease. Clin. Microbiol. Rev. *10*, 320–344.

Gumpert, J., Todorov, T., and Toshkov, A. (1980). Light microscopic investigations on lysozyme-and penicillin-induced morphological changes in Erysipelothrix rhusiopathiae and on propagation of its protoplast type L-form. Z. Allg. Mikrobiol. 20, 431–440.

Innes, C.M.J., and Allan, E.J. (2001). Induction, growth and antibiotic production of Streptomyces viridifaciens L-form bacteria. J. Appl. Microbiol. 90, 301–308.

Kawai, Y., Mercier, R., Wu, L.J., Domínguez-Cuevas, P., Oshima, T., and Errington, J. (2015). Cell growth of wall-free L-form bacteria is limited by oxidative damage. Curr. Biol. *25*, 1613–1618.

Kawai, Y., Mickiewicz, K., and Errington, J. (2018). Lysozyme Counteracts β-Lactam Antibiotics by Promoting the Emergence of L-Form Bacteria. Cell *172*, this issue, 1038–1049.

Klieneberger, E. (1935). The natural occurrence of pleuropneumonia-like organism in apparent symbiosis with *Strrptobacillus moniliformis* and other bacteria. J. Pathol. Bacteriol. *40*, 93–105.

Owens, W.E. (1987). Isolation of Staphylococcus aureus L forms from experimentally induced bovine mastitis. J. Clin. Microbiol. *25*, 1956–1961.

Learning from Everyday Images Enables Expert-like Diagnosis of Retinal Diseases

Ladislav Rampasek^{1,2,3} and Anna Goldenberg^{1,2,3,*}

¹University of Toronto, Department of Computer Science, Toronto, ON, Canada

²The Hospital for Sick Children, Toronto, ON, Canada

³Vector Institute, Toronto, ON, Canada

*Correspondence: anna.goldenberg@utoronto.ca https://doi.org/10.1016/j.cell.2018.02.013

Kermany et al. report an application of a neural network trained on millions of everyday images to a database of thousands of retinal tomography images that they gathered and expert labeled, resulting in a rapid and accurate diagnosis of retinal diseases.

Recent years have seen considerable increase in the development of machinelearning algorithms for medical imaging to support clinical decision making and diagnostics. Advancements in medical imaging and visualization provide medical experts with powerful tools to gain and analyze visual representation of the body. Thanks to such insights, better and faster diagnosis or triage is made possible, providing doctors with extra time to, for example, deal with difficult cases. Computed tomography, magnetic resonance imaging, ultrasound, or radiology generate high-resolution digital imaging data. Increased automatization and availability of these technologies have contributed to a great increase in the volume of imaging data, making the medical imaging field poised to

greatly benefit from machine-learning approaches. In this issue of *Cell*, Kermany et al. (2018) show how a computer vision system for classification of natural images can be successfully adapted to diagnosing common blinding retinal diseases and pediatric pneumonia, achieving performance comparable to that of human experts.

Computer processing and elements of computer vision have been used in medical imaging since the 1990s. Advanced processing and analytics help to augment work processes of human experts, providing them with image-derived scores that augment their analysis and increase their efficiency. Today, machine learning is the prime tool ready to greatly improve these workflows and perhaps completely automate diagnostic

and referral tests in the future. Kermany et al. (2018) provide a glimpse of this future. In their study, Kermany et al. demonstrate the use of deep learning for triage and diagnosis of choroidal neovascularization, diabetic macular edema, and drusen, three common retinal diseases. Authors have retrospectively analyzed optical coherence tomography (OCT) images of retina from the Shiley Eye Institute of the University of California San Diego and the Shanghai First People's Hospital, eventually yielding 108,312 OCT images curated by human experts. Leveraging these data, their deep-learning model in the form of a convolutional neural network outperformed two out of six ophthalmologists using a combined measure of sensitivity and precision. Interestingly, when their neural



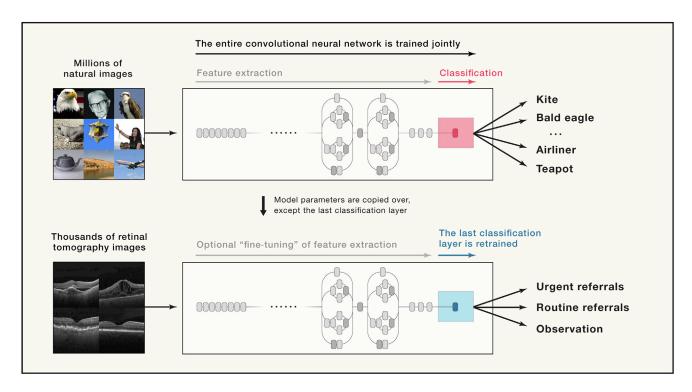


Figure 1. Schematic of a Convolutional Neural Network and Transfer Learning

Generalizable feature extraction is learned on a loosely related set of tasks but with much larger datasets, such as ImageNet, with 1.2 million images across 1,000 categories. The number of images and variety of classes in the dataset are important for the learning of reusable features. The task-specific data are used to train task-specific upper layers of the neural network and optionally to adjust the copied-over part of the model. The depicted CNN architecture is a simplification of Google's Inception-v3, Szegedy et al., 2016, where each rectangular component represents a computational layer, e.g., a convolutional or pooling layer.

network was trained on only 4,000 (~25 times fewer) labeled images, the error increased only 2-fold compared to the full dataset. To a great extent, these results are due to the transfer learning approach, where a more general computer vision model that categorizes common objects in natural images is adapted to a new, more specific, task by transferring some of the learned parameters to the new domain. Kermany et al. (2018) further demonstrate the generality of this approach by applying it to establish a model for the diagnosis of viral and bacterial pediatric pneumonia using 5,232 chest X-ray images.

Transfer learning became a popular approach to classifying medical images due to the relatively small number of medical images compared to the total number of natural images that can be used to train a general model. In the past 2–3 years, several papers have applied models originally trained on natural images to medical imaging in radiography, pathology, and cell imaging, starting with the work of Bar et al. (2015) for detection of different

types of chest pathologies in X-ray images. Soon followed the works of Gulshan et al. (2016) in detection of diabetic retinopathy in retinal fundus photographs, Esteva et al. (2017), who demonstrated dermatologist-level classification of skin cancer, and many others (for an extensive survey, see Litjens et al., 2017).

Neural networks, such as convolutional neural nets used by Kermany et al., process the raw input pixel data of images/ videos in a hierarchical fashion by feeding the data through multiple subsequent computational layers that compute increasingly complex, yet distilled, representations of the original input. The last layer of the network can predict the desired output (e.g., identify what is in the image) from the final representation provided by the penultimate layer. All of these layers are trained on large datasets, with the most popular being ImageNet (Deng et al., 2009), the basis for the annual ImageNet Large-Scale Visual Recognition Challenge (2010-2017). This competition, aimed at developing the best vision models for natural image classification,

saw the rise of convolutional neural networks, from the pioneering AlexNet through VGG, ResNet, GoogLeNet, and Inception, to name a few. Particularly, the Inception-v3 network by Szegedy et al., 2016—trained on ImageNet—has become the favorite choice for transfer learning in the past two years (Gulshan et al., 2016; Esteva et al., 2017, Kermany et al., 2018; and others).

In the process of training these networks to classify millions of images, it has been shown by many that the early layers learn to extract basic image features, such as edges, which are used by subsequent layers to recognize more complex features, like corners, parallel edges, or other general object parts. In fact, the neural network's ability to learn such effective feature extraction is what made deep learning (meaning stacking many layers in a deep architecture) so successful, greatly improving over hand-crafted feature extraction designed by human experts. It was then noted (Yosinski et al., 2014) that such a feature-extraction "pipeline" learned by

a neural network on one big dataset is often robust enough to transfer well to other tasks and datasets. This technique of transfer learning has enabled successful applications of deep-learning models on tasks with limited training data availability, as is the case in the medical field. The generalizable feature-extracting part of the network is borrowed from a model trained on a large number of loosely related or unrelated images, while the task-specific data are used to train task-specific upper layers of the network, such as the last classifier (Figure 1). At the same time, it is also possible to "fine-tune" the borrowed layers, although in the case of Kermany et al. (2018), it is interesting to note that this did not improve performance.

As this study and the papers released on this topic in the last 2-3 years show, the field of machine learning for medical imaging is rapidly expanding and holds great potential to bring better and potentially faster and more affordable care to patients. While the transfer learning approach has allowed to kickstart this field even in cases where relatively limited training data are available, many questions remain: for example, is it possible to improve the accuracy from more specific background training? One could imagine the development of an ImageNet-like database for medical images across diseases, which might improve the recognition of more specific features. How can networks be adapted to analyze intrinsically different images, such as 3D scans? What information is needed to implement some of these solutions in other hospitals in more remote or third-world countries to properly account for the inevitable bias in these new datasets?

REFERENCES

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., and Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature 542, 115-118.

Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316, 2402-2410

Kermany, D.S., Goldbaum, M., Cai, W., Valentim, C.C.S., Liang, H., Baxter, S.L., McKeown, A.,

Yang, G., Wu, X., Yan, F., et al. (2018). Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. Cell 172, this issue, 1122-1131.

Bar, Y., Diamant, I., Wolf, L., Lieberman, S., Konen, E., and Greenspan, H. (2015). Chest pathology detection using deep learning with non-medical training. In 2015 IEEE 12th International Symposium, pp. 294-297.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B., and Sánchez, C.I. (2017). A Survey on Deep Learning in Medical Image Analysis. arXiv, arXiv:1702.05747, https:// arxiv.org/abs/1702.05747.

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

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision. In 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2818-2826.

Yosinski, J., Clune, J., Bengio, Y., and Lipson, H. (2014). How transferable are features in deep neural networks? In NIPS'14 Proceedings of the 27th International Conference on Neural Information Processing Systems 2, 3320-3328.