**Criminal identification and Tendency for feature reconstruction to exhibit a bias for one shot image classification using Siamese Neural Networks and GAN**

**Abstract**

In this paper, we propose a novel approach that addresses the challenge of learning good features in machine learning applications, especially in scenarios with limited data availability. where accurate predictions need to be made based on just a single example per new class. To tackle this, we employ a Siamese network architecture consisting of two identical network and Face generation using GAN. By calculating the absolute difference between the feature vectors extracted from the two images and computing the similarity score, we enable the network to learn discriminative representations. In contrast, humans demonstrate a remarkable ability to quickly understand and recognize new patterns, even with minimal exposure, facilitating future recognition of variations on these concepts. While machine learning has achieved impressive performance in various applications, it often struggles when making predictions about data with limited supervised information. Our approach mitigates this issue by limiting assumptions about input structure and automatically acquiring features that promote successful generalization, leveraging the use of different activation functions for criminal identification. Better result when we compare with existing model.

INTROUDCTION

Siamese nets were first introduced in the early 1990s by Bromley and LeCun to solve signature verification as an image matching problem (Bromley et al., 1993). A Siamese neural network consists of twin networks which accept distinct inputs but are joined by an energy function at the top. This function computes some metric between the highest level feature representation on each side.. The parameters between the twin networks are tied. Weight tying guarantees that two extremely similar images could not possibly be mapped by their respective networks to very different locations in feature space because each network computes the same function. Also, the network is symmetric, so that whenever we present two distinct images to the twin networks, the top conjoining layer will compute the same metric as if we were to we present the same two images but to the opposite twins. In LeCun et al., the authors used a contrastive energy function which contained dual terms to decrease the energy of like pairs and increase the energy of unlike pairs . However, in this paper we use the weighted L1 distance between the twin feature vectors h1 and h2 combined with a sigmoid activation, which maps onto the interval [0; 1]. Thus a cross-entropy objective is a natural choice for training the network. Note that in LeCun et al., they directly learned the similarity metric, which was implicitly defined. only those output units which were the result of complete overlap between each convolutional filter and the input feature maps. The units in the final convolutional layer are flattened into a single vector. This convolutional layer is followed by a fully-connected layer, and then one more layer computing the induced distance metric between each Siamese twin, which is given to a single sigmoidal output unit. More precisely, the prediction vector , where is the sigmoidal activation function. This final layer induces a metric on the learned feature space of the (L1)th hidden layer and scores the similarity between the two feature vectors. parameters that are learned by the model during training, weighting the importance of the component-wise distance. This defines a final Lully-connected layer for the network which joins the two siamese twins. We depict one example above , which shows the largest version of our model that we considered. This network also gave the best result for any network on the verification task. In recent years, the field of machine learning has been revolutionized by the advent of Generative Adversarial Networks (GANs), a groundbreaking framework that has demonstrated remarkable capabilities in generating realistic and diverse data. GANs have found applications across a wide spectrum of domains, including image synthesis, text generation, music composition, and even drug discovery. This paper provides an insightful introduction to the fundamental concepts, architecture, and applications of GANs in contemporary research. By exploring the underlying principles that enable GANs to produce data indistinguishable from real samples, this paper aims to equip researchers with a foundational understanding of this powerful technique and its potential to drive innovation in various fields.

Approch

In general, we learn image representations via a supervised metric-based approach with siamese neural networks, then reuse that network’s features for one-shot learning without any retraining. In our experiments, we restrict our attention to character recognition, although the basic approach can be replicated for almost any modality . For this domain, we employ large Siamese convolutional neural networks which are capable of learning generic image features useful for making predictions about unknown class distributions even when very few examples from these new distributions are available are easily trained using standard optimization techniques on pairs sampled from the source data and provide a competitive approach that does not rely upon domain-specific knowledge by instead exploiting deep learning techniques. To develop a model for one-shot image classification, we aim to first learn a neural network that can discriminate between the class-identity of image pairs, which is the standard verification task for image recognition. We hypothesize that networks which do well at at verification. should generalize to one-shot classification. The verification model learns to identify input pairs according to the probability that they belong to the same class or different classes. This model can then be used to evaluate new images, exactly one per novel class, in a pairwise manner against the test image. The pairing with the highest score according to the verification network is then awarded the highest probability for the one-shot task. If the features learned by the verification model are sufficient to confirm or deny the identity of images from one set of images, then they ought to be sufficient for other images, provided that the model has been exposed to a variety of images to encourage variance . The heart of a GAN consists of two main components—the generator and the discriminator—engaged in a dynamic interplay. The generator aims to synthesize data instances from random noise, while the discriminator's task is to differentiate between real data and generated data. As training progresses, the generator strives to improve its output to deceive the discriminator, while the discriminator refines its ability to distinguish real from generated samples. This iterative process drives both networks to improve their performance, ultimately converging towards a point where the generated data becomes increasingly difficult to distinguish from real data.

**Literature Review**

AlexeyDosovitskiy, LucasBeyer, AlexanderKolesnikov(2021). While the Transformer architecture has become the de-facto standard for natural language processing tasks, its applications to computer vision remain limited. In vision, attention is either applied in conjunction with convolutional networks, or used to replace certain components of convolutional networks while keeping their overall structure in place. We show that this reliance on CNNs is not necessary and a pure transformer applied directly to sequences of image patches can perform very well on image classification task . It give the accuracy of 87% and build the main research gap from this paper is require to have large amount of dataset and high computation

1. Chen, W., Gong, X., & Wang, Z. (2021) Neural Architecture Search (NAS) has been explosively studied to automate the discovery of top-performer neural networks. Current works require heavy training of supernet or intensive architecture evaluations, thus suffering from heavy resource consumption and often incurring search bias due to truncated training or approximations. Can we select the best neural architectures without involving any training and eliminate a drastic portion of the search cost? We provide an affirmative answer, by proposing a novel framework called training-free neural architecture search (TE-NAS). TE-NAS ranks architectures by analyzing the spectrum of the neural tangent kernel (NTK) and the number of linear regions in the input space. Both are motivated by recent theory advances in deep networks and can be computed without any training and any label. We show that: (1) these two measurements imply the trainability and expressivity of a neural network; (2) they strongly correlate with the network's test accuracy. Further on, we design a pruning-based NAS mechanism to achieve a more flexible and superior trade-off between the trainability and expressivity during the search. In NAS-Bench-201 and DARTS search spaces, TE-NAS completes high-quality search but only costs 0.5 and 4 GPU hours with one 1080Ti on CIFAR-10 and ImageNet, respectively. We hope our work inspires more attempts in bridging the theoretical findings of deep networks and practical impacts in real NAS applications
2. Lu, D., & Weng, Q. (2007) Image classification is a complex process that may be affected by many factors. This paper examines current practices, problems, and prospects of image classification. The emphasis is placed on the summarization of major advanced classification approaches and the techniques used for improving classification accuracy. In addition, some important issues affecting classification performance are discussed. This literature review suggests that designing a suitable image‐processing procedure is a prerequisite for a successful classification of remotely sensed data into a thematic map. Effective use of multiple features of remotely sensed data and the selection of a suitable classification method are especially significant for improving classification accuracy. Non‐parametric classifiers such as neural network, decision tree classifier, and knowledge‐based classification have increasingly become important approaches for multisource data classification. Integration of remote sensing, geographical information systems (GIS), and expert system emerges as a new research frontier. More research, however, is needed to identify and reduce uncertainties in the image‐processing chain to improve classification accuracy.
3. Keysers, D., Deselaers, T., Gollan, C., & Ney, H. (2007).We present the application of different nonlinear image deformation models to the task of image recognition. The deformation models are especially suited for local changes as they often occur in the presence of image object variability. We show that, among the discussed models, there is one approach that combines simplicity of implementation, low-computational complexity, and highly competitive performance across various real-world image recognition tasks. We show experimentally that the model performs very well for four different handwritten digit recognition tasks and for the classification of medical images, thus showing high generalization capacity. In particular, an error rate of 0.54 percent on the MNIST benchmark is achieved, as well as the lowest reported error rate, specifically 12.6 percent, in the 2005 international ImageCLEF evaluation of medical image specifically categorization.
4. Qiu, J., Zhang, Z. G., Xie, Z., & Lin, S. (2019) The convolution layer has been the dominant feature extractor in computer vision for years. However, the spatial aggregation in convolution is basically a pattern matching process that applies fixed filters which are inefficient at modeling visual elements with varying spatial distributions. This paper presents a new image feature extractor, called the local relation layer, that adaptively determines aggregation weights based on the compositional relationship of local pixel pairs. With this relational approach, it can composite visual elements into higher-level entities in a more efficient manner that benefits semantic inference. A network built with local relation layers, called the Local Relation Network (LR-Net), is found to provide greater modeling capacity than its counterpart built with regular convolution on large-scale recognition tasks such as ImageNet classification.
5. He, K., Zhang, X., Ren, S., & Sun, J. (2016) Deeper neural networks are more difficult to train. We present a residual learning framework to ease the training of networks that are substantially deeper than those used previously. We explicitly reformulate the layers as learning residual functions with reference to the layer inputs, instead of learning unreferenced functions. We provide comprehensive empirical evidence showing that these residual networks are easier to optimize, and can gain accuracy from considerably increased depth. On the ImageNet dataset we evaluate residual nets with a depth of up to 152 layers---8x deeper than VGG nets but still having lower complexity. An ensemble of these residual nets achieves 3.57% error on the ImageNet test set. This result won the 1st place on the ILSVRC 2015 classification task. We also present analysis on CIFAR-10 with 100 and 1000 layers. The depth of representations is of central importance for many visual recognition tasks. Solely due to our extremely deep representations, we obtain a 28% relative improvement on the COCO object detection dataset. Deep residual nets are foundations of our submissions to ILSVRC & COCO 2015 competitions, where we also won the 1st places on the tasks of ImageNet detection, ImageNet localization, COCO detection, and COCO segmentation.
6. Van Der Walt, S., Schönberger, J. L & Yu, T. ET al (2014) scikit-image is an image processing library that implements algorithms and utilities for use in research, education and industry applications. It is released under the liberal Modified BSD open source license, provides a well-documented API in the Python programming language, and is developed by an active, international team of collaborators. In this paper we highlight the advantages of open source to achieve the goals of the scikit-image library, and we showcase several real-world image processing applications that use scikit-image. More information can be found on the project homepage
7. Lu, D., & Weng, Q. (2007b)Image classification is a complex process that may be affected by many factors. This paper examines current practices, problems, and prospects of image classification. The emphasis is placed on the summarization of major advanced classification approaches and the techniques used for improving classification accuracy. In addition, some important issues affecting classification performance are discussed. This literature review suggests that designing a suitable image‐processing procedure is a prerequisite for a successful classification of remotely sensed data into a thematic map. Effective use of multiple features of remotely sensed data and the selection of a suitable classification method are especially significant for improving classification accuracy. Non‐parametric classifiers such as neural network, decision tree classifier, and knowledge‐based classification have increasingly become important approaches for multisource data classification. Integration of remote sensing, geographical information systems (GIS), and expert system emerges as a new research frontier. More research, however, is needed to identify and reduce uncertainties in the image‐processing chain to improve classification accuracy.
8. Zhao, H., Jia, J., & Koltun, V. (2020)Recent work has shown that self-attention can serve as a basic building block for image recognition models. We explore variations of self-attention and assess their effectiveness for image recognition. We consider two forms of self-attention. One is pairwise self-attention, which generalizes standard dot-product attention and is fundamentally a set operator. The other is patchwise self-attention, which is strictly more powerful than convolution. Our pairwise self-attention networks match or outperform their convolutional counterparts, and the patchwise models substantially outperform the convolutional baselines. We also conduct experiments that probe the robustness of learned representations and conclude that self-attention networks may have significant benefits in terms of robustness and generalization
9. Egmont-Petersen, M., De Ridder, D., & Handels, H. (2002)We review more than 200 applications of neural networks in [image processing](https://www.sciencedirect.com/topics/computer-science/image-processing) and discuss the present and possible future role of neural networks, especially feed-forward neural networks, Kohonen feature maps and [Hopfield neural networks](https://www.sciencedirect.com/topics/computer-science/hopfield-neural-network). The various applications are categorised into a novel two-dimensional taxonomy for [image processing algorithms](https://www.sciencedirect.com/topics/computer-science/image-processing-algorithm). One dimension specifies the type of task performed by the algorithm: preprocessing, data reduction/feature extraction, segmentation, object recognition, image understanding and optimisation. The other dimension captures the abstraction level of the input data processed by the algorithm: pixel-level, local feature-level, structure-level, object-level, object-set-level and scene characterisation. Each of the six types of tasks poses specific constraints to a neural-based approach. These specific conditions are discussed in detail. A synthesis is made of unresolved problems related to the application of pattern recognition techniques in image processing and specifically to the application of neural networks. Finally, we present an outlook into the future application of neural networks and relate them to novel developments.
10. Pang, L., Lan, Y., Guo, J., Xu, J., Wan, S., & Cheng, X. (2016).Matching two texts is a fundamental problem in many natural language processing tasks. An effective way is to extract meaningful matching patterns from words, phrases, and sentences to produce the matching score. Inspired by the success of convolutional neural network in image recognition, where neurons can capture many complicated patterns based on the extracted elementary visual patterns such as oriented edges and corners, we propose to model text matching as the problem of image recognition. Firstly, a matching matrix whose entries represent the similarities between words is constructed and viewed as an image. Then a convolutional neural network is utilized to capture rich matching patterns in a layer-by-layer way. We show that by resembling the compositional hierarchies of patterns in image recognition, our model can successfully identify salient signals such as n-gram and n-term matchings. Experimental results demonstrate its superiority against the baselines.
11. State-of-the-art image classification methods require an intensive learning/training stage (using SVM, Boosting, etc.) In contrast, non-parametric nearest-neighbor (NN) based image classifiers require no training time and have other favorable properties. However, the large performance gap between these two families of approaches rendered NN-based image classifiers useless. We claim that the effectiveness of non-parametric NN-based image classification has been considerably undervalued. We argue that two practices commonly used in image classification methods, have led to the inferior performance of NN-based image classifiers: (i) Quantization of local image descriptors (used to generate "bags-of-words ", codebooks). (ii) Computation of 'image-to-image' distance, instead of 'image-to-class' distance. We propose a trivial NN-based classifier - NBNN, (Naive-Bayes nearest-neighbor), which employs NN- distances in the space of the local image descriptors (and not in the space of images). NBNN computes direct 'image- to-class' distances without descriptor quantization. We further show that under the Naive-Bayes assumption, the theoretically optimal image classifier can be accurately approximated by NBNN. Although NBNN is extremely simple, efficient, and requires no learning/training phase, its performance ranks among the top leading learning-based image classifiers. Empirical comparisons are shown on several challenging databases (Caltech-101 ,Caltech-256 and Graz-01).
12. Tokozume, Y., Ushiku, Y., & Harada, T. (2018). In this paper, we propose a novel learning method for image classification called Between-Class learning (BC learning). We generate between-class images by mixing two images belonging to different classes with a random ratio. We then input the mixed image to the model and train the model to output the mixing ratio. BC learning has the ability to impose constraints on the shape of the feature distributions, and thus the generalization ability is improved. BC learning is originally a method developed for sounds, which can be digitally mixed. Mixing two image data does not appear to make sense; however, we argue that because convolutional neural networks have an aspect of treating input data as waveforms, what works on sounds must also work on images. First, we propose a simple mixing method using internal divisions, which surprisingly proves to significantly improve performance. Second, we propose a mixing method that treats the images as waveforms, which leads to a further improvement in performance. As a result, we achieved 19.4% and 2.26% top-1 errors on ImageNet-1K and CIFAR-10, respectively.
13. Fujiyoshi, H., Hirakawa, T., & Yamashita, T. (2019). Various image recognition tasks were handled in the image recognition field prior to 2010 by combining image local features manually designed by researchers (called handcrafted features) and machine learning method. After entering the 2010, However, many image recognition methods that use deep learning have been proposed. The image recognition methods using deep learning are far superior to the methods used prior to the appearance of deep learning in general object recognition competitions. Hence, this paper will explain how deep learning is applied to the field of image recognition, and will also explain the latest trends of deep learning-based [autonomous driving](https://www.sciencedirect.com/topics/social-sciences/autonomous-driving).
14. Xie, C., Tan, M., Gong, B., Wang, J., Yuille, A. L., & Le, Q. V. (2020). Adversarial examples are commonly viewed as a threat to ConvNets. Here we present an opposite perspective: adversarial examples can be used to improve image recognition models if harnessed in the right manner. We propose AdvProp, an enhanced adversarial training scheme which treats adversarial examples as additional examples, to prevent overfitting. Key to our method is the usage of a separate auxiliary batch norm for adversarial examples, as they have different underlying distributions to normal examples. We show that AdvProp improves a wide range of models on various image recognition tasks and performs better when the models are bigger. For instance, by applying AdvProp to the latest EfficientNet-B7 [28] on ImageNet, we achieve significant improvements on ImageNet (+0.7%), ImageNet-C (+6.5%), ImageNet-A (+7.0%), Stylized-ImageNet (+4.8%). With an enhanced EfficientNet-B8, our method achieves the state-of-the-art 85.5% ImageNet top-1 accuracy without extra data. This result even surpasses the best model in [20] which is trained with 3.5B Instagram images ( 3000X more than ImageNet) and 9.4X more parameters. Code and models will be made publicly available.
15. Gao, S., Tsang, I. W., & Chia, L. (2010). Recent research has shown the effectiveness of using sparse coding(Sc) to solve many computer vision problems. Motivated by the fact that kernel trick can capture the nonlinear similarity of features, which may reduce the feature quantization error and boost the sparse coding performance, we propose Kernel Sparse Representation(KSR). KSR is essentially the sparse coding technique in a high dimensional feature space mapped by implicit mapping function. We apply KSR to both image classification and face recognition. By incorporating KSR into Spatial Pyramid Matching(SPM), we propose KSRSPM for image classification. KSRSPM can further reduce the information loss in feature quantization step compared with Spatial Pyramid Matching using Sparse Coding(ScSPM). KSRSPM can be both regarded as the generalization of Efficient Match Kernel(EMK) and an extension of ScSPM. Compared with sparse coding, KSR can learn more discriminative sparse codes for face recognition. Extensive experimental results show that KSR outperforms sparse coding and EMK, and achieves state-of-the-art performance for image classification and face recognition on publicly available datasets.
16. Wang, C., Blei, D. M., & Li, F. (2009). Image classification and annotation are important problems in computer vision, but rarely considered together. Intuitively, annotations provide evidence for the class label, and the class label provides evidence for annotations. For example, an image of class highway is more likely annotated with words “road,” “car,” and “traffic” than words “fish,” “boat,” and “scuba.” In this paper, we develop a new probabilistic model for jointly modeling the image, its class label, and its annotations. Our model treats the class label as a global description of the image, and treats annotation terms as local descriptions of parts of the image. Its underlying probabilistic assumptions naturally integrate these two sources of information. We derive an approximate inference and estimation algorithms based on variational methods, as well as efficient approximations for classifying and annotating new images. We examine the performance of our model on two real-world image data sets, illustrating that a single model provides competitive annotation performance, and superior classification performance.
17. Chen, C., Li, O., Tao, D., Barnett, A., Rudin, C., & Su, J. (2019). When we are faced with challenging image classification tasks, we often explain our reasoning by dissecting the image, and pointing out prototypical aspects of one class or another. The mounting evidence for each of the classes helps us make our final decision. In this work, we introduce a deep network architecture -- prototypical part network (ProtoPNet), that reasons in a similar way: the network dissects the image by finding prototypical parts, and combines evidence from the prototypes to make a final classification. The model thus reasons in a way that is qualitatively similar to the way ornithologists, physicians, and others would explain to people on how to solve challenging image classification tasks. The network uses only image-level labels for training without any annotations for parts of images. We demonstrate our method on the CUB-200-2011 dataset and the Stanford Cars dataset. Our experiments show that ProtoPNet can achieve comparable accuracy with its analogous non-interpretable counterpart, and when several ProtoPNets are combined into a larger network, it can achieve an accuracy that is on par with some of the best-performing deep models. Moreover, ProtoPNet provides a level of interpretability that is absent in other interpretable deep models.
18. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). The latest generation of convolutional neural networks (CNNs) has achieved impressive results in the field of image classification. This paper is concerned with a new approach to the development of plant disease recognition model, based on leaf image classification, by the use of deep convolutional networks. Novel way of training and the methodology used facilitate a quick and easy system implementation in practice. The developed model is able to recognize 13 different types of plant diseases out of healthy leaves, with the ability to distinguish plant leaves from their surroundings. According to our knowledge, this method for plant disease recognition has been proposed for the first time. All essential steps required for implementing this disease recognition model are fully described throughout the paper, starting from gathering images in order to create a database, assessed by agricultural experts. Caffe, a deep learning framework developed by Berkley Vision and Learning Centre, was used to perform the deep CNN training. The experimental results on the developed model achieved precision between 91% and 98%, for separate class tests, on average 96.3%.
19. Dollár, P., Tu, Z., Tao, H., & Belongie, S. (2007). The efficiency and robustness of a vision system is often largely determined by the quality of the image features available to it. In data mining, one typically works with immense volumes of raw data, which demands effective algorithms to explore the data space. In analogy to data mining, the space of meaningful features for image analysis is also quite vast. Recently, the challenges associated with these problem areas have become more tractable through progress made in machine learning and concerted research effort in manual feature design by domain experts. In this paper, we propose a feature mining paradigm for image classification and examine several feature mining strategies. We also derive a principled approach for dealing with features with varying computational demands. Our goal is to alleviate the burden of manual feature design, which is a key problem in computer vision and machine learning. We include an in-depth empirical study on three typical data sets and offer theoretical explanations for the performance of various feature mining strategies. As a final confirmation of our ideas, we show results of a system, that utilizing feature mining strategies matches or outperforms the best reported results on pedestrian classification (where considerable effort has been devoted to expert feature design).
20. Xie, S., Kirillov, A. M., Girshick, R., & He, K. (2019). Neural networks for image recognition have evolved through extensive manual design from simple chain-like models to structures with multiple wiring paths. The success of ResNets and DenseNets is due in large part to their innovative wiring plans. Now, neural architecture search (NAS) studies are exploring the joint optimization of wiring and operation types, however, the space of possible wirings is constrained and still driven by manual design despite being searched. In this paper, we explore a more diverse set of connectivity patterns through the lens of randomly wired neural networks. To do this, we first define the concept of a stochastic network generator that encapsulates the entire network generation process. Encapsulation provides a unified view of NAS and randomly wired networks. Then, we use three classical random graph models to generate randomly wired graphs for networks. The results are surprising: several variants of these random generators yield network instances that have competitive accuracy on the ImageNet benchmark. These results suggest that new efforts focusing on designing better network generators may lead to new breakthroughs by exploring less constrained search spaces with more room for novel design.
21. He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., & Li, M. (2019). Much of the recent progress made in image classification research can be credited to training procedure refinements, such as changes in data augmentations and optimization methods. In the literature, however, most refinements are either briefly mentioned as implementation details or only visible in source code. In this paper, we will examine a collection of such refinements and empirically evaluate their impact on the final model accuracy through ablation study. We will show that, by combining these refinements together, we are able to improve various CNN models significantly. For example, we raise ResNet-50's top-1 validation accuracy from 75.3% to 79.29% on ImageNet. We will also demonstrate that improvement on image classification accuracy leads to better transfer learning performance in other application domains such as object detection and semantic segmentation.
22. Traore, B. B., Kamsu-Foguem, B., & Tangara, F. (2018).During an epidemic crisis, medical image analysis namely microscopic analyses are made to confirm or not the existence of the epidemic [pathogen](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/pathogen) in suspected cases. Pathogen are all infectious agents such as a virus, bacterium, [protozoa](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/protozoa), prion etc. However, there is often a lack of specialists in the handling of microscopes, hence allowing the need to make the microscopic analysis abroad. This results in a considerable loss of time and in the meantime, the epidemic continues to spread. To save time in the analysis of samples, we propose to make the future microscopes more intelligent so that they will be able to indicate by themselves the existence or not of the [pathogen](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/pathogen) of an epidemic in a sample. To have a smart microscope, we propose a methodology based on efficient Convolution [Neural Network](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/neural-network) (CNN) architecture in order to classify epidemic pathogen with five deep learning phases: (1) Training dataset of provided images (2) CNN Training (3) Testing data preparation (4) CNN generated model on testing data and finally (5) Evaluation of images classified. The resulted classification process can be integrated in a mobile computing solution on future microscopes. CNN can improve the accuracy in pathogens diagnosis that are focused on hand-tuned feature extraction implying some human mistakes. For our study, we consider cholera and malaria epidemics for microscopic [images classification](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/image-classification) with a relevant CNN, respectively [Vibrio cholerae](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/vibrio-cholerae) images and [Plasmodium falciparum](https://www.sciencedirect.com/topics/agricultural-and-biological-sciences/plasmodium-falciparum) images. Image classification is the task of taking an input image and outputting a class or a probability of classes that best describes the image. Interesting results have been obtained from the CNN model generated achieving the classification accuracy of 94%, with 200 Vibrio cholera images and 200 *Plasmodium falciparum* images for training dataset and 80 images for testing data. Although this document addresses the classification of epidemic pathogen images using a CNN model, the underlying principles apply to the other fields of [science and technology](https://www.sciencedirect.com/topics/earth-and-planetary-sciences/science-and-technology), because of its performance and its capability to handle more layers than the previous traditional neural networks.
23. Ciresan, D., Meier, U., & Schmidhuber, J. (2012). Traditional methods of computer vision and machine learning cannot match human performance on tasks such as the recognition of handwritten digits or traffic signs. Our biologically plausible, wide and deep artificial neural network architectures can. Small (often minimal) receptive fields of convolutional winner-take-all neurons yield large network depth, resulting in roughly as many sparsely connected neural layers as found in mammals between retina and visual cortex. Only winner neurons are trained. Several deep neural columns become experts on inputs preprocessed in different ways; their predictions are averaged. Graphics cards allow for fast training. On the very competitive MNIST handwriting benchmark, our method is the first to achieve near-human performance. On a traffic sign recognition benchmark it outperforms humans by a factor of two. We also improve the state-of-the-art on a plethora of common image classification benchmarks
24. Qiao, S., Liu, C., Shen, W., & Yuille, A. L. (2018). In this paper, we are interested in the few-shot learning problem. In particular, we focus on a challenging scenario where the number of categories is large and the number of examples per novel category is very limited, e.g. 1, 2, or 3. Motivated by the close relationship between the parameters and the activations in a neural network associated with the same category, we propose a novel method that can adapt a pre-trained neural network to novel categories by directly predicting the parameters from the activations. Zero training is required in adaptation to novel categories, and fast inference is realized by a single forward pass. We evaluate our method by doing few-shot image recognition on the ImageNet dataset, which achieves the state-of-the-art classification accuracy on novel categories by a significant margin while keeping comparable performance on the large-scale categories. We also test our method on the MiniImageNet dataset and it strongly outperforms the previous state-of-the-art methods.

References

1 Dosovitskiy, A., Beyer, L., Kolesnikov, A. I., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2021). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. In *International Conference on Learning Representations*. <https://openreview.net/pdf?id=YicbFdNTTy>

3. Lu, D., & Weng, Q. (2007). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, *28*(5), 823–870. https://doi.org/10.1080/01431160600746456

4. Chen, W., Gong, X., & Wang, Z. (2021). Neural Architecture Search on ImageNet in Four GPU Hours: A Theoretically Inspired Perspective. In *International Conference on Learning Representations*. <https://openreview.net/pdf?id=Cnon5ezMHtu>

5. Qiu, J., Zhang, Z. G., Xie, Z., & Lin, S. (2019). *Local Relation Networks for Image Recognition*. <https://doi.org/10.1109/iccv.2019.00356>

6. He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep Residual Learning for Image Recognition*. <https://doi.org/10.1109/cvpr.2016.90>

7. Van Der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., & Yu, T. (2014). scikit-image: image processing in Python. *PeerJ*, *2*, e453. <https://doi.org/10.7717/peerj.453>

8. Lu, D., & Weng, Q. (2007b). A survey of image classification methods and techniques for improving classification performance. *International Journal of Remote Sensing*, *28*(5), 823–870. https://doi.org/10.1080/01431160600746456

9. Zhao, H., Jia, J., & Koltun, V. (2020). *Exploring Self-Attention for Image Recognition*. <https://doi.org/10.1109/cvpr42600.2020.01009>

10. Egmont-Petersen, M., De Ridder, D., & Handels, H. (2002). Image processing with neural networks—a review. *Pattern Recognition*, *35*(10), 2279–2301. https://doi.org/10.1016/s0031-3203(01)00178-9

11. Pang, L., Lan, Y., Guo, J., Xu, J., Wan, S., & Cheng, X. (2016). Text Matching as Image Recognition. *AAAI Conference*, *30*(1). <https://doi.org/10.1609/aaai.v30i1.10341>

12. Boiman, O., Shechtman, E., & Irani, M. (2008). *In defense of Nearest-Neighbor based image classification*. https://doi.org/10.1109/cvpr.2008.4587598

13. Tokozume, Y., Ushiku, Y., & Harada, T. (2018). *Between-Class Learning for Image Classification*. https://doi.org/10.1109/cvpr.2018.00575

14. Fujiyoshi, H., Hirakawa, T., & Yamashita, T. (2019). Deep learning-based image recognition for autonomous driving. *Iatss Research*, *43*(4), 244–252. https://doi.org/10.1016/j.iatssr.2019.11.008

15. Xie, C., Tan, M., Gong, B., Wang, J., Yuille, A. L., & Le, Q. V. (2020). *Adversarial Examples Improve Image Recognition*. <https://doi.org/10.1109/cvpr42600.2020.00090>

16. Gao, S., Tsang, I. W., & Chia, L. (2010). Kernel Sparse Representation for Image Classification and Face Recognition. In *Springer eBooks* (pp. 1–14). https://doi.org/10.1007/978-3-642-15561-1\_1

17. Wang, C., Blei, D. M., & Li, F. (2009). Simultaneous image classification and annotation. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*. <https://doi.org/10.1109/cvpr.2009.5206800>

18. Chen, C., Li, O., Tao, D., Barnett, A., Rudin, C., & Su, J. (2019). This Looks Like That: Deep Learning for Interpretable Image Recognition. In *Neural Information Processing Systems* (Vol. 32, pp. 8928–8939). <https://papers.nips.cc/paper/9095-this-looks-like-that-deep-learning-for-interpretable-image-recognition.pdf>

19. Sladojevic, S., Arsenovic, M., Anderla, A., Culibrk, D., & Stefanovic, D. (2016). Deep Neural Networks Based Recognition of Plant Diseases by Leaf Image Classification. *Computational Intelligence and Neuroscience*, *2016*, 1–11. <https://doi.org/10.1155/2016/3289801>

20. Dollár, P., Tu, Z., Tao, H., & Belongie, S. (2007). *Feature Mining for Image Classification*. <https://doi.org/10.1109/cvpr.2007.383046>

21. Xie, S., Kirillov, A. M., Girshick, R., & He, K. (2019). *Exploring Randomly Wired Neural Networks for Image Recognition*. <https://doi.org/10.1109/iccv.2019.00137>

22. He, T., Zhang, Z., Zhang, H., Zhang, Z., Xie, J., & Li, M. (2019). *Bag of Tricks for Image Classification with Convolutional Neural Networks*. <https://doi.org/10.1109/cvpr.2019.00065>

23. Traore, B. B., Kamsu-Foguem, B., & Tangara, F. (2018). Deep convolution neural network for image recognition. *Ecological Informatics*, *48*, 257–268. <https://doi.org/10.1016/j.ecoinf.2018.10.002>

24. Ciresan, D., Meier, U., & Schmidhuber, J. (2012). *Multi-column deep neural networks for image classification*. <https://doi.org/10.1109/cvpr.2012.6248110>

25. Qiao, S., Liu, C., Shen, W., & Yuille, A. L. (2018). *Few-Shot Image Recognition by Predicting Parameters from Activations*. https://doi.org/10.1109/cvpr.2018.00755