Linking Research to The Real world

Anurag Malyala
2K15/CO/035

Department of Computer Science Engineering

Seminar Report CO-403

Table of Contents

Introduction 1
Background 2
GPUs and High-Performance Computing 2
Artificial Intelligence vs Machine Learning vs Deep Learning 4
Artificial Intelligence 4
Machine Learning 4
Deep Learning 5
Convolutional Neural Networks 5
Recurrent Neural Networks. 7
Applications 8
Health Care 8
Applying Deep Learning to Metastatic Breast Cancer Detection 8
Automating Drug Discovery 9
Image Processing 11
Learning to See in the Dark 11
Image Restoration and Denoising 13
Conclusion 17
References 18

Delhi Technological University

Abstract

Department of Computer science and Engineering

Bachelor of Technology (CO-403)

Linking Research to The Real world

By Anurag Malyala 2K15/CO/035

The real world and the research fields live a world of disconnect in most aspects. This however is not necessarily the case. The increase in Computer Science research activity as seen a correlation in real world as well. This report looks into the driving factors behind this surge in research and its trickling down into the real world by studying the mathematical and algorithmic concepts and the hardware technologies that enable them to be distributed to their destination markets. We also look into two use cases where research has made its way into products and services in Medicine and Image processing by exploring their working and significance.

Introduction

Over the past few years the number of research publications in computer science have sky rocketed. One of the main drivers of this boom are the breakthroughs in High Performance Computing by the means of GPU (Graphical Processing Units) providing cheap and accessible parallel processing capabilities previously only dreamt of. The field that has seen the highest traction is Artificial Intelligence, especially the sub domains for Machine Learning and Deep Learning. But all these breakthroughs and research is of no value if they're not applied to solving real world problems through innovative products and service. This report covers four such products and services, covering of range of research areas, Computer Vision and Graphics and AI Enabled Computational Medicare

It can be seen from the aforementioned fields there exists a common theme, all these falls under the umbrella of Artificial Intelligence, this can be attributed to the fact that these are "Hard" problems, that is, as of now, there is no way of solving them in polynomial time. Here, the contribution of Artificial Intelligence comes into play. AI provides a set of methods and techniques for solving such problem through search and approximation.

Combined with techniques in signal processing, these AI methods have surpassed conventional methods by a large factor. Where previously, the task of object detection and recognition was a heavily involved process requiring manual selection of key points and parameters for classifying, to give results which could only be considered satisfactory. Today, we have methods which require no manual selection processes and give near human level results (Christian Szegedy, 2016), that too at considerably lower computational Costs.

Same is the case in Natural Language tasks, where once, real-time translation (Ron J. Weiss, 2017) between any pair of languages would've been considered near impossible if not fiction has become a basic feature available on most smartphones. Even in healthcare, the latest works from Academia and Industry (Liu, 2018) (Steiner, et al., 2018) have reached accuracy levels only matched by humans, and in some cases even better.

The rest of the report is divided into sections describing the background work that enables these technologies, shared methodologies and finally goes into describing a select few examples of products and services which have made their way from papers and journals into our daily life.

Background

This section covers the background concepts driving the research Engine, first we look at GPUs and HPC which are enabling the fast pace of research and then the main Techniques and Algorithms relevant to the services covered.

GPUs and High-Performance Computing

At its core, all AI methods can be decomposed into operations on vectors and matrixes. Thus, allowing the use of highly efficient algorithms from Linear Algebra. Consider images as in example, digital images are represented as a grid of values, called pixels, and depending on the image type this can be an 8-bit image with values between 0 and 255 or 10-bit images with values from 0 to 1024. Even text data can be represented as vectors and matrixes, either in terms of term-frequencies or word vectors.

The algorithms involved in processing these matrices and vectors perform repeated matrix multiplications, in fact, the most popular method for classification and object recognitions, Convolutional Neural Networks are defined as matrix multiplication operations.

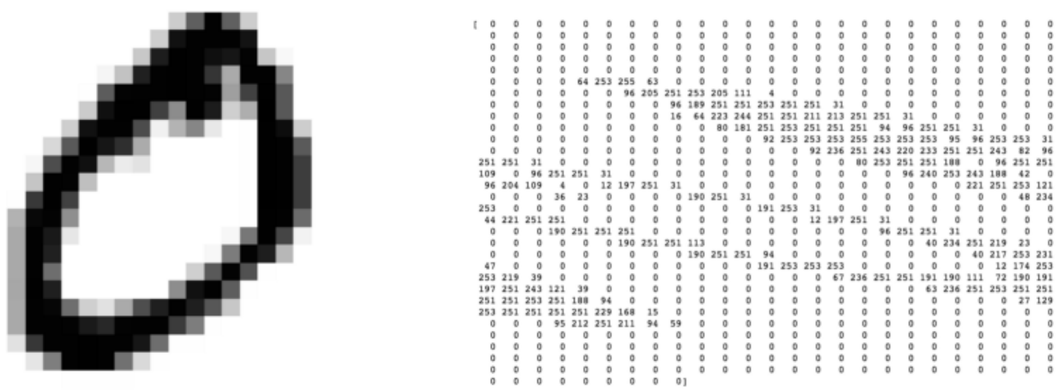


Figure 1 An image represented as a grid of numbers

But why do we need HPC and GPUs? This can be answered from the fact that matrix multiplication operation is extremely parallelable, the standard matrix multiplication operation is an order of N^3 operation, while, given adequate parallel processors, this can be brought down to a logarithmic operation. CPUs have what we call generalized processor cores, these are extremely powerful processing units designed to do many operations with vary low latencies, and in any given commercial CPU package we can find at most 8 cores, with server and workstation grade CPUs going up to 44 cores in extreme cases. This is also the key differentiating factor between CPU and GPUs.

Figure 2 Difference between CPU and GPU architecture

GPUs have a high count of densely packed specialized processor cores, with even entry level processors have core counts in the hundreds, most of which are designed to do only one operation, but at high bandwidths churning through data order of magnitudes faster than CPUs.

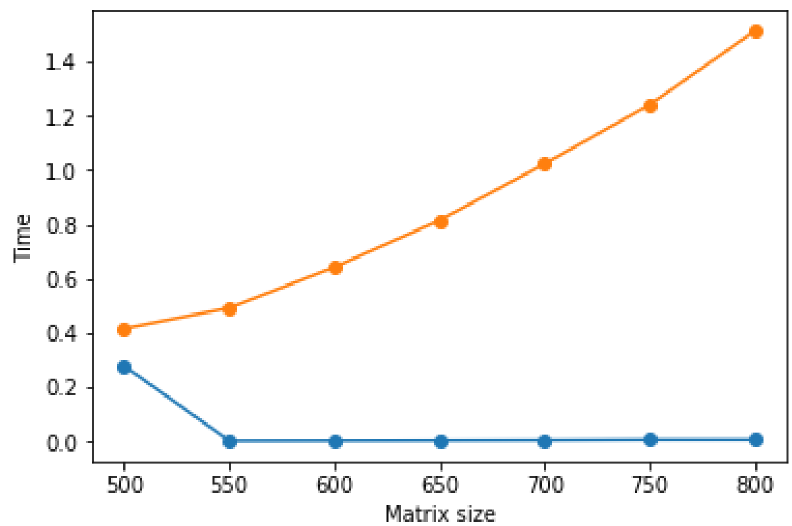
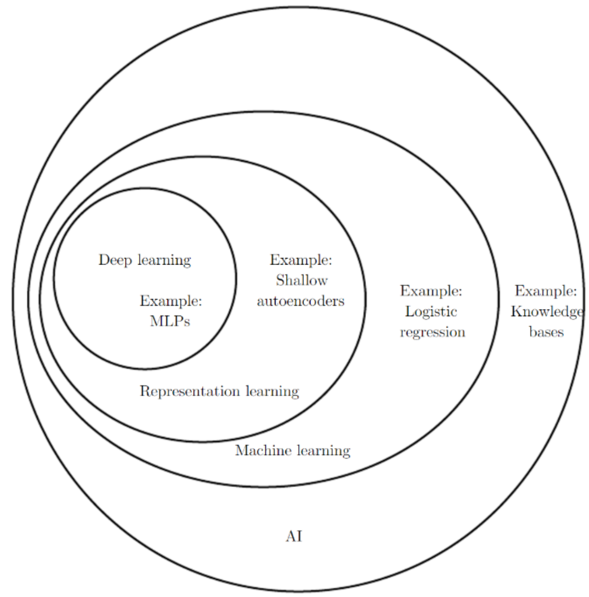


Figure 3 CPU vs GPU, Increase in computation time with increase in input size of matrix. Comparing Nvidia GTX 1080 with latest Generation Inter Core i7

Manufacturers like Nvidia have risen to the demand of such high core count, high memory processing units and have started developing products specialized for this task, the Tesla V100 being one of the best if not The Best processor available in market.

Artificial Intelligence vs Machine Learning vs Deep Learning

Artificial Intelligence



As the name suggests, artificial intelligence can be loosely interpreted to mean incorporating human intelligence to machines. Artificial intelligence is the broader concept that consists of everything from Good Old-Fashioned AI (GOFAI) all the way to futuristic technologies such as deep learning. Whenever a machine completes tasks based on a set of stipulated rules that solve problems (algorithms), such an "intelligent" behavior is what is called artificial intelligence.

AI-powered machines are usually classified into two groups—general and narrow. The general artificial intelligence AI machines can intelligently solve problems, like the ones mentioned above. The narrow intelligence AI machines can perform specific tasks very well, sometimes better than humans—though they are limited in scope.

Machine Learning

A sub-field of AI, whose focus is to use data to train computer algorithms to perform tasks that typically cannot be done (or very difficult to accomplish) through hard wiring the logic into a program, because no one is quite sure what the rules are. Recognizing objects in images is one such task where the rules are unclear.

As the name suggests, machine learning can be loosely interpreted to mean empowering computer systems with the ability to "learn". The intention of ML is to enable machines to learn by themselves using the provided data and make accurate predictions. ML is a subset of artificial intelligence; in fact, it's simply a technique for realizing AI. It is a method of training algorithms such that they can learn how to make decisions. Training in machine learning entails giving a lot of data to the algorithm and allowing it to learn more about the processed information.

Deep Learning

Deep learning is a subset of machine learning. Deep artificial neural networks are a set of algorithms reaching new levels of accuracy for many important problems, such as image recognition, sound recognition, recommender systems, etc.

It uses machine learning techniques to solve real-world problems by tapping into neural networks that simulate human decision-making. Deep learning can be costly and requires huge datasets to train itself. This is because there are a huge number of parameters that need to be understood by a learning algorithm, which can primarily yield a lot of false-positives. For example, a deep learning algorithm could be trained to 'learn' how a dog looks like. It would take an enormous dataset of images for it to understand the minor details that distinguish a dog from a wolf or a fox.

Convolutional Neural Networks

Now that we're done with definitions, we can look into the two key breakthroughs in the AI, first Convolutional Neural Networks and then Recurrent Neural Networks.

Convolutional Neural networks (Y. Lecun, 1998) allow computers to see, in other words, Convnets are used to recognize images by transforming the original image through layers to a class score. CNN was inspired by the visual cortex. Every time we see something, a series of layers of neurons gets activated, and each layer will detect a set of features such as lines, edges. The high level of layers will detect more complex features in order to recognize what we saw.

A CNN has Two Parts, a feature learner, and a classifier. The input to them are RGB or Black and White Images, usually of the format $[HxWxC]$ where C is 3 in case of RGB and 1 in case of black and white.

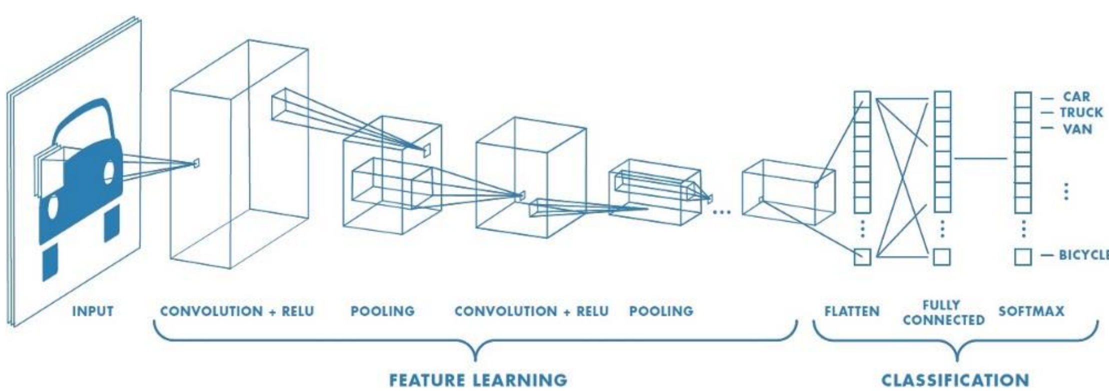


Figure 4 The two components of a CNN

The feature learning portion consists of filters, which try to learn to extract features from the input. These filters move across the image processing a section of image at a time and producing an output for the next layer to process.

Filter, Kernel, or Feature Detector: is a small matrix used for features detection.

Feature Map: is the output formed by sliding the filter over the image and computing the dot product.

Receptive field: is a local region of the input volume that has the same size as the filter.

Between two layers, there exists a Non-Linearity which promotes learning of non-linear functions to capture the information stored in the images. Figure 4 shows a ReLU (Nair, 2010), i.e. a Rectified Linear Unit, given by $f(x) = \max(0, x)$

The pooling Function exists to 1. Reduce the parameters which allows processing large images. 2. Retains the "activity" from the last layer this not losing any information between layers.

Finally, the classifier is used to give the "class scores", the classifier can be any classifier model but, in most cases, a fully connected neural network is used for this task.

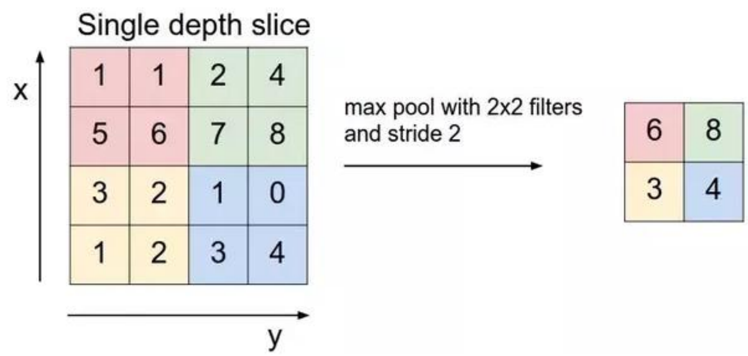


Figure 5 The pooling Operation

The entire CNN can be summarized as follows:

- Provide input image into convolution layer
- Choose parameters, apply filters with strides, padding if requires. Perform convolution on the image and apply ReLU activation to the matrix.
- Perform pooling to reduce dimensionality size
- Add as many convolutional layers until satisfied
- Flatten the output and feed into a fully connected layer (FC Layer)
- Output the class using an activation function (Logistic Regression with cost functions) and classifies images

Recurrent Neural Networks.

As per Wikipedia, a recurrent neural network (RNN) is a class of artificial neural network where connections between units form a directed graph along a sequence. This allows it to exhibit dynamic temporal behavior for a time sequence. Unlike feedforward neural networks, RNNs can use their internal state (memory) to process sequences of inputs. This makes them applicable to tasks such as unsegmented, connected handwriting recognition or speech recognition.

RNN remembers what it knows from previous input using a simple loop. This loop takes the information from previous time stamp and adds it to the input of current time stamp.

Figure 6 shows the basic RNN structure. X_t is the input to the network and h_t is the output of the network at time "t". A is an RNN cell. RNN cells contain neural networks just like a feed-forward network or a perceptron. The hidden states of the feed-forward network are again used along with the input in next step.

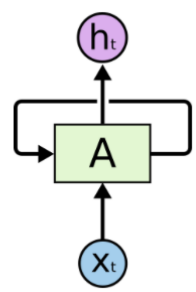


Figure 6 A recurrent Node (Olah, 2015)

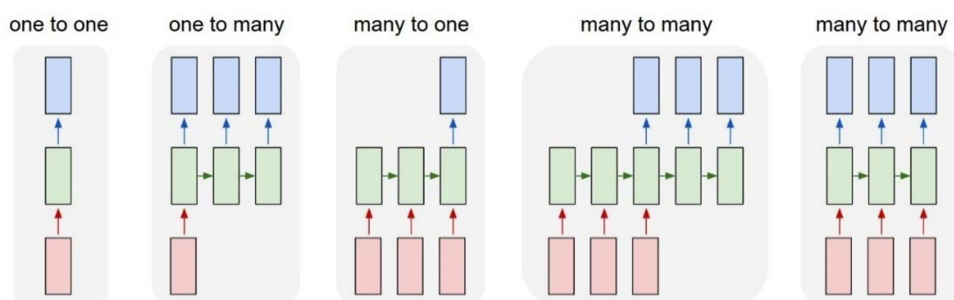


Figure 7 Each rectangle is a vector and arrows represent functions (e.g. matrix multiply). Input vectors are in red, output vectors are in blue and green vectors hold the RNN's state (more on this soon). From left to right: (1) Vanilla mode of processing without RNN, from fixed-sized input to fixed-sized output (e.g. image classification). (2) Sequence output. (3) Sequence input. (4) Sequence input and sequence output (5) Synced sequence input and output. (Karpathy, 2015)

The semantic information of the sequence is preserved in the hidden states of the recurrent neural network. This semantic information keeps on altering as the new input is observed and is again passed to the next input. This leads to the seamless flow of sequential information through the network of RNN. Passing of information from one time to other helps in finding the correlation among the events or words in a sentence and is often known as "long-term dependencies".

Applications

Health Care

Applying Deep Learning to Metastatic Breast Cancer Detection

In October 2018, Google published the results of their work in diagnosing Metastic Breast Cancer, claiming an accuracy of 99%. Such a high accuracy is in fact better than the current gold standard, a pathologist with a microscope. In 2017, Google introduced LYNA which provided gigapixel-sized pathology slides of lymph nodes from breast cancer patients for researchers to develop computer algorithms to detect metastatic cancer. In the work, "Detecting Cancer Metastases on Gigapixel Pathology Images" (Yun Liu, 2017) presented a convolutional neural network (CNN) approach for segmenting gigapixel pathology images into normal and cancerous pixels to aid breast cancer diagnosis.

They presented an automated approach to detect and localize tumors as small as 100×100 pixels in digitized microscopy images sized 100, 000×100, 000 pixels or larger. leveraging the Inception (V3) neural network architecture, a type of CNN, and obtains state-of-the-art results on the Camelyon16 dataset in the challenging lesion-level tumor detection task. They obtained a score of 8 false positives per image detect 92.4% of the tumors, relative to 82.7% by the previous best automated approach. For comparison, a human pathologist achieved 73.2% sensitivity using exhaustive search.

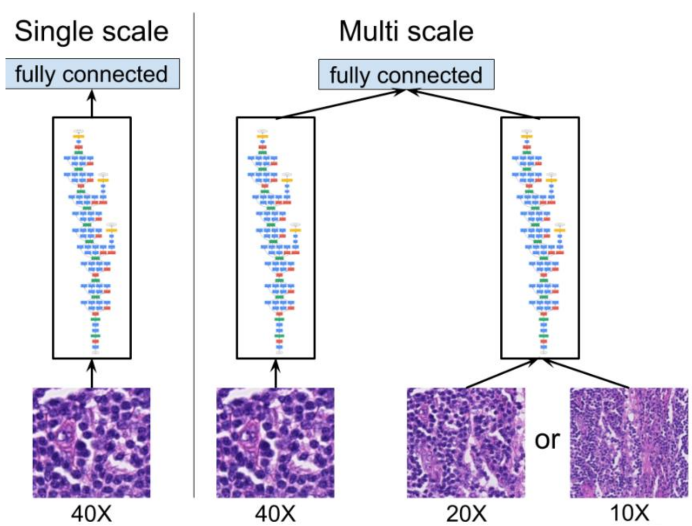


Figure 8 The three colorful blocks represent Inception (V3) towers up to the second-last layer (PreLogit). Single scale utilizes one tower with input images at 40X magnification; multi-scale utilizes multiple (e.g.,2) input magnifications that are input to separate towers and merged.

They proposed the following methodology:

For each input image, they chose to process the image in patches of 299 by 299 for each image the classified them as containing a cancerous tumor or a normal cell around a central region of 128 by 128. They label a patch as cancerous if at least one of the pixels in the center regions is labeled cancerous. They also explored the influence of the number of parameters by reducing the number of filters while keeping the number of layers constant

Because of the large bias between cancerous and normal patches where is the patches containing tumors or of the order of 0.01% to 70% depending on the size of the patch. They applied several data augmentations to combat this bias by including rotations and flips.

Finally, they performed by sliding through the scans in steps of 128 to match the center pixels and then applying the augmentations, using them to do the actual classification.

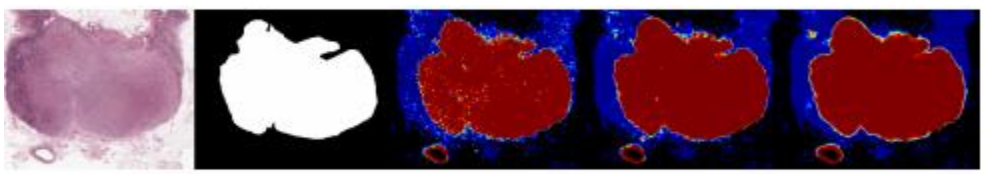


Figure 9 Left to right: sample image, ground truth (tumor in white), and heatmap outputs (40X-ensemble-of-3, 40X+20X, and 40X+10X). Heatmaps of 40X and 40Xensemble-of-3 look identical. The red circular regions at the bottom left quadrant of the heatmaps is unannotated tumor. Some of the speckles are either out of focus patches on the image or non-tumor patches within a large tumor.

Automating Drug Discovery

Protein crystallization is a key step to biomedical research concerned with discovering the structure of complex biomolecules. Because that structure determines the molecule's function, it helps scientists design new drugs that are specifically targeted to that function. However, protein crystals are rare and difficult to find. Hundreds of experiments are typically run for each protein, and while the setup and imaging are mostly automated, finding individual protein crystals remains largely performed through visual inspection and thus prone to human error. Critically, missing these structures can result in lost opportunity for important biomedical discoveries for advancing the state of medicine.

In their work "Classification of crystallization outcomes using deep convolutional neural networks" (Andrew E. Bruno, 2018) used some of the most recent architectures of deep convolutional networks and customized them to achieve an accuracy of more than 94% on the visual recognition task of identifying protein crystals.

Due to the large variability between imaging technologies and data acquisition approaches, coming up with a single approach to the visual recognition problem may appear daunting. Crystals can be very small, which makes them rare structures in a large image containing otherwise undifferentiated visual clutter.

Using the Inception model, the authors developed a model to identify proteins to a degree of precision and speed that it is usable for automatic drug discovery.

The model gives a 4 class output as shown in Figure 10. The 4 classes correspond to a distillation of 38 sub classes. This is done to reduce the variance in data between protein types. To get back the particular protein types furether classification is needed.

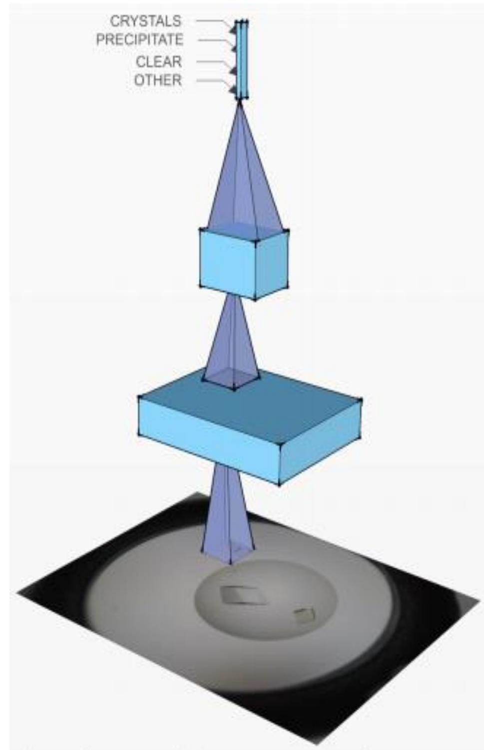


Figure 10 Conceptual representation of the model used

They modified the network to use high resolution images to account for the fine nature of the crystalline structures. The generic nature of the model used by them shows that it can be extended beyond just classification to other human-mediated visual recognition task in drug discovery.

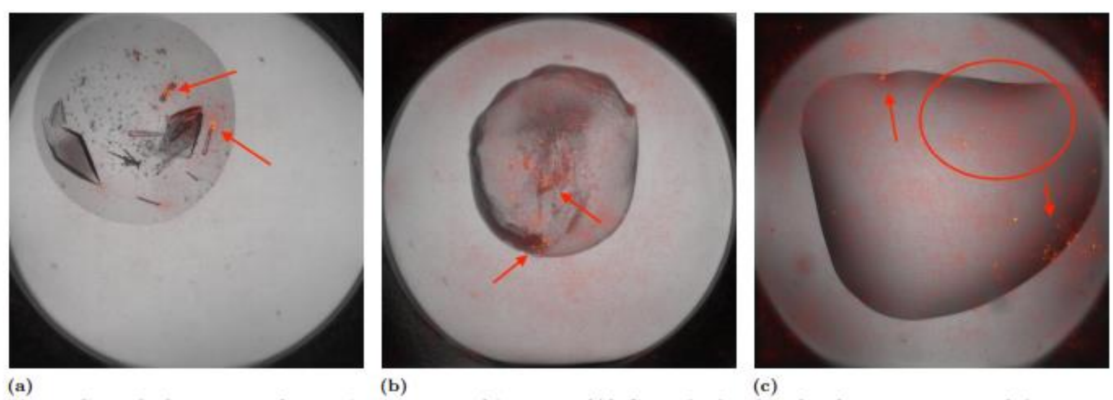


Figure 11 (A) Crystal: the classifier focuses on some of the angular geometric features of individual crystals (arrows). (B) Precipitate: the classifier lands on the precipitate (arrows). (C) Clear: The classifier broadly samples the image, likely because this label is characterized by the absence of structures rather than their presence

Image Processing

Learning to See in the Dark

In their latest flagship, Google introduced a feature called Night Sight, a new feature of the Pixel Camera app that lets you take sharp, clean photographs in very low light, even in light so dim you can't see much with your own eyes.

Low light photography is hard, anybody who has photographed a dimly lit scene will be familiar with image noise, which looks like random variations in brightness from pixel to pixel. For smartphone cameras, which have small lenses and sensors, a major source of noise is the natural variation of the number of photons entering the lens, called shot noise. Every camera suffers from it, and it would be present even if the sensor electronics were perfect. However, they are not, so a second source of noise are random errors introduced when converting the electronic charge resulting from light hitting each pixel to a number, called read noise. These and other sources of randomness contribute to the overall signal-to-noise ratio (SNR), a measure of how much the image stands out from these variations in brightness. Fortunately, SNR rises with the square root of exposure time (or faster), so taking a longer exposure produces a cleaner picture. But it’s hard to hold still long enough to take a good picture in dim light, and whatever you're photographing probably won't hold still either.

Now, being google, we’ll probably never know the technology they’ve used in their phones, but in the paper “Learning to See in the Dark” (Chen Chen, 2018) the authors introduced a method that uses the raw output from cameras and can generate images with ample lighting from images which to the untrained eye might look pitch black.

(a) Camera output with ISO 8,000 (b) Camera output with ISO 409,600 (c) Our result from the raw data of (a)

Figure 12 Results of the proposed algorithm

The traditional pipeline for taking low light photos involves steps like, white balance, demosaicing, denoising, sharpening, color space conversion, gamma correction, and others. These modules are often tuned for specific cameras. Jiang et al. (H. Jiang, 2017) proposed to use a large collection of local, linear, and learned (L3) filters to approximate the complex nonlinear pipelines found in modern consumer imaging systems. Yet neither the traditional pipeline nor the L3 pipeline successfully deal with fast low-light imaging, as they are not able to handle the extremely low SNR. Similarly, another proposed method (S. W. Hasinoff, 2016) used a burst imaging pipeline for smartphone cameras. This method can produce good results by aligning and blending multiple images, but introduces a certain level of complexity, for example due to the need for dense correspondence estimation, and may not easily extend to video capture, for example due to the use of lucky imaging.

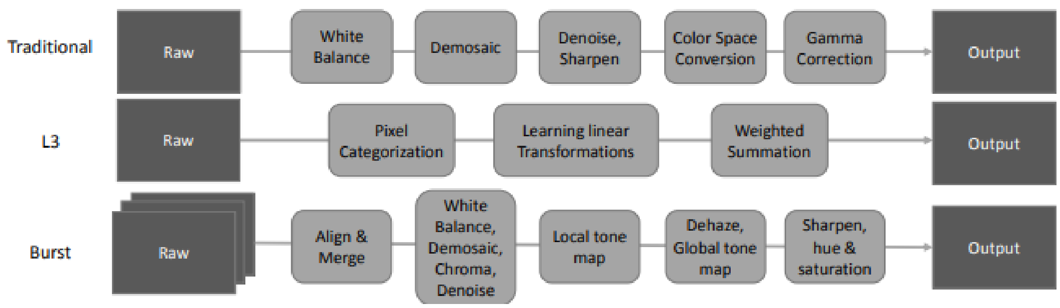


Figure 13 Traditional Methods

In their work, they proposed an end to end, model that does in real time processing of low-light images using the raw output from the sensor.

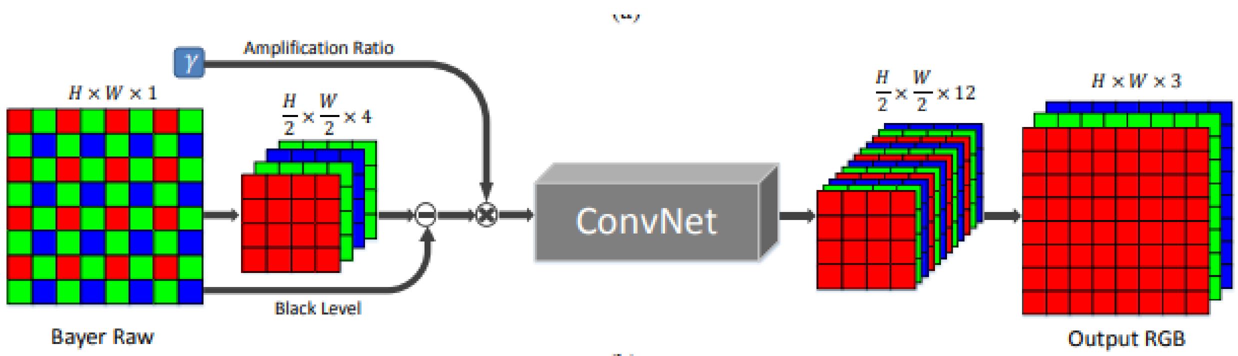


Figure 14 The proposed pipeline

From the Bayer arrays, they packed the input into four channels and correspondingly reduce the spatial resolution by a factor of two in each dimension. For X-Trans arrays the raw data is arranged in 6×6 blocks; which are packed into 9 channels instead of 36 channels by exchanging adjacent elements. They then subtract the black level and scale the data by the desired amplification ratio (e.g., x100 or x300). The packed and amplified data is fed into a fully-convolutional network. The output is a 12-channel image with half the spatial resolution. This half-sized output is processed by a sub-pixel layer to recover the original resolution

The CNN used by the authors follows the U-Net architecture (O. Ronneberger, 2015). To control the final brightness the amplification ratio is externally supplied to allow the user to adjust it.

It is evident from the results that the proposed system works perfectly.

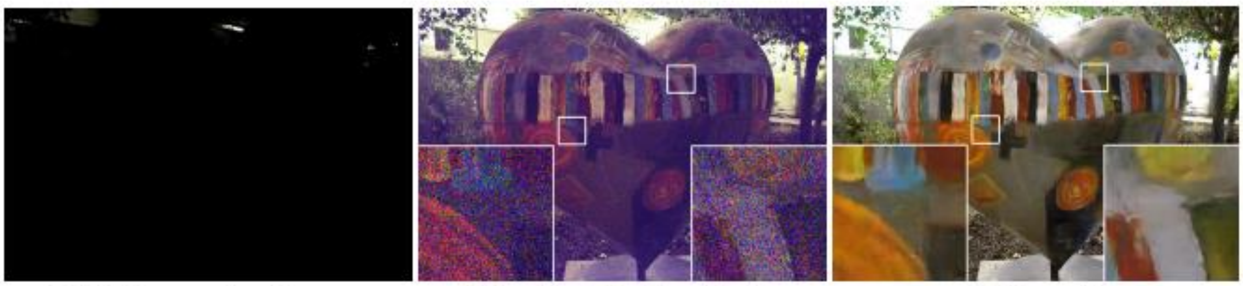


Figure 15 (a) An image captured at night by the Fujifilm X-T2 camera with ISO 800, aperture f/7.1, and exposure of 1/30 second. The illuminance at the camera is approximately 1 lux. (b) Processing the raw data by a traditional pipeline does not effectively handle the noise and color bias in the data. (c) Our result obtained from the same raw data (Chen Chen, 2018)

Image Restoration and Denoising

The task of restoring and recovering old photos of which no digital, if not physical copies exist is a labor-intensive task often taking months to finish a single piece. Both Nvidia and Adobe in their recent publications have presented methods for denoising, de-artifacting and filling in images in a matter of seconds.

Image Denoising

Fixing grainy or noisy images is a difficult task, especially when the availability of clean ground truth images is impossible. This task not only has applications in Photography but more profound ones in astronomical analysis, medical imaging, and animation.

In the work "Noise2Noise: Learning Image Restoration without Clean Data" (Jaakko Lehtinen, 2018) The authors learn to learn the noise present in the images. They find that it is in fact not necessary to have clear images and in fact sometimes it is better not to include them.

They use an encoder-decoder style CNN for learning the noise representation. Compared to other proposed methods, their method uses a very simple model thus allowing very fast inference. As a consequence of learning the noise model, this method can also remove text overlays from images as well.

The implications of such denoising and restoring capabilities are huge. In medical imaging, denoising MRI scans is an important task, the proposed method gives a high degree of denoising even when the input MRI Scans are extremely noisy.

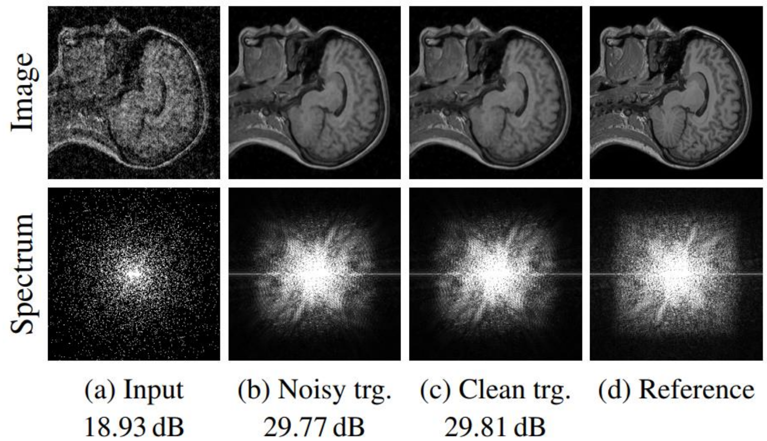


Figure 16 MRI reconstruction example. (a) Input image with only 10% of spectrum samples retained and scaled by 1/p. (b) Reconstruction by a network trained with noisy target images similar to the input image. (c) Same as previous, but training done with clean target images similar to the reference image. (d) Original, uncorrupted image. PSNR values refer to the images shown here, see text for averages over the entire validation set. (Jaakko Lehtinen, 2018)

Another industry that has benefited from this is the computer animation industry. Physically accurate renderings of virtual environments are most often generated through a process known as Monte Carlo path tracing. This amounts to drawing random sequences of scattering events (“light paths”) in the scene that connect light sources and virtual sensors, and integrating the radiance carried by them over all possible paths. The Monte Carlo integrator is constructed such that the intensity of each pixel is the expectation of the random path sampling process, i.e., the sampling noise is zero-mean. However, despite decades of research into importance sampling techniques, little else can be said about the distribution. It varies from pixel to pixel, heavily depends on the scene configuration and rendering parameters, and can be arbitrarily multimodal. Some lighting effects, such as focused caustics, also result in extremely long-tailed distributions with rare, bright outliers.

To get clean images the simulations have to run for hours if not days, by having such a denoising step the rendering times can get cut drastically bringing down costs of production.

Figure 17 Denoising a Monte Carlo rendered image. (a) Image rendered with 64 samples per pixel. (b) Denoised 64 spp input, trained using 64 spp targets. (c) Same as previous but trained on clean targets. (d) Reference image rendered with 131 072 samples per pixel. PSNR values refer to the images shown here, see text for averages over the entire validation set.

Image Restoration

Restoring images which may have gotten sections torn off or ruined because of some accidents is an important task in restoring antiques. In the Paper "Image Inpainting for Irregular Holes Using Partial Convolutions" (Guilin Liu, 2018), propose the use of partial convolutions, where the convolution is masked and renormalized to be conditioned on only valid pixels.

Their approach not only allows recovering missing sections but also allows removing objects from the image, used extensively in image retouching. In fact, this feature in some form has existed in Adobe's Photoshop software since version 6 (CS6).

The proposed model uses stacked partial convolution operations and mask updating steps to perform image inpainting. A partial convolution only updates the values of the output if the region on which it is processing contains pixels, i.e. the section is not missing or has been marked for removal.

They also use a U-Net like architecture for their model. But replace all convolutional layers with partial ones and perform up sampling for the decoder. They do not use any existing padding schemes for our convolutions when near image boundaries. Instead, the partial convolution layer directly handles this with appropriate masking. This will further ensure that the inpainted content at the image border will not be affected by invalid values outside of the image, which can be interpreted as another hole.

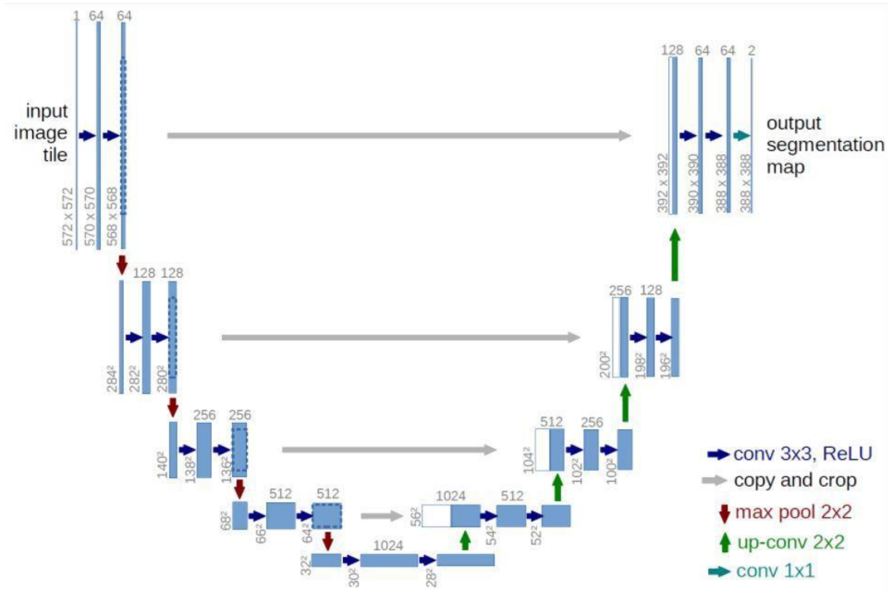


Figure 18 The encoder-decoder U-Net model used by the authors

While traditional methods approximated the missing pixels from the surrounding ones, the proposed model learns the local reference by building an intuition of the regions. This also allows the model to handle cases with huge cut-outs in images.

The model to some extent can handle image super-resolution that is, upscaling images synthetically. To achieve this the authors, introduce holes in the image by off-setting images pixels by one and then performing in filling operations. This is repeated till the desired resolution is achieved.

(a) Input (b) PM (c) GL (d) GntIpt (e) PConv (f) GT

Figure 19 Test results on the standard ImageNet dataset

Conclusion

From the above examples we can see that correlation between research and real-world use cases exist and have huge implications on day to day activities. It may not be apparent at first but as is the nature of research, the applications of them come under niches too. Most MRI processing machines a Denoising model. Nearly all image rendering suites, like Cycles (Foundation, n.d.), Vray (Group, n.d.), and Renderman (Pixar, n.d.) have a denoising model built in to cut render times for low power machines.

Image processing software's like Photoshop have got built in features for in-filling and restoration. Software like Adobe Lightroom has a feature that can take raw images and generate bright images from images otherwise unusable.

Google's Breast cancer detection model has become the go to tool for Pathologists for screening through cancer tests, thus reducing the time for diagnosis drastically which can potentially save lives.

Other Domains where research has trickled down to trivial products are personal assistants like Amazon Alexa and Google Assistant.

References

Andrew E. Bruno, P. C. (2018). Classification of crystallization outcomes using deep convolutional neural networks. *PLOS ONE*.

Chen Chen, Q. C. (2018). Learning to See in the Dark. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR).*

Christian Szegedy, S. I. (2016). Inception-v4, Inception-ResNet and the Impact of Residual Connections on learning. *CoRR*. Retrieved from http://arxiv.org/abs/1602.07261

Foundation, B. (n.d.). *Cycles*. Retrieved from Blender: https://www.cycles-renderer.org/

Group, C. (n.d.). *V-ray*. Retrieved from chaos group: www.chaosgroup.com

Guilin Liu, F. A.-C. (2018). "Image Inpainting for Irregular Holes Using Partial Convolutions". *ECCV.*

H. Jiang, Q. T. (2017). Learning the image processing pipeline. *IEEE Transactions on Image Processing.* IEEE.

Jaakko Lehtinen, J. M. (2018). Noise2Noise: Learning Image Restoration without Clean Data. *International Conference on Machine Learning.*

Karpathy, A. (2015, 05 21). *Andrej Karpathy blog*. Retrieved from GitHub: https://karpathy.github.io/2015/05/21/rnn-effectiveness/

Liu, Y. a. (2018). Artificial Intelligence–Based Breast Cancer Nodal Metastasis Detection. *Archives of Pathology & Laboratory Medicine*.

Nair, V. a. (2010). Rectified Linear Units Improve Restricted Boltzmann Machines. *Proceedings of the 27th International Conference on International Conference on Machine Learning*, (pp. 807--814). Haifa, Israel.

O. Ronneberger, P. F. (2015). U-net: Convolutional networks for biomedical image segmentation. *MIC-CAI*.

Olah, C. (2015, 08 27). *Understanding LSTM Networks*. Retrieved from github: https://colah.github.io/posts/2015-08-Understanding-LSTMs/

Pixar. (n.d.). *Renderman*. Retrieved from pixar: https://renderman.pixar.com/

Ron J. Weiss, J. C. (2017). Sequence-to-Sequence Models Can Directly Translate Foreign Speech. *Interspeech.*

S. W. Hasinoff, D. S. (2016). Burst photography for high dynamic range and low-light imaging on mobile cameras. *ACM Transactions on Graphics, 35(6).*

Steiner, D. F., MacDonald, R., Liu, Y., Truszkowski, P., Hipp, J. D., Gammage, C., . . . Stumpe, M. C. (2018). Impact of Deep Learning Assistance on the Histopathologic Review of Lymph Nodes for Metastatic Breast Cancer. *The American Journal of Surgical Pathology*.

Y. Lecun, L. B. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE, vol. 86, no. 11*, (pp. 2278-2324).

Yun Liu, K. G. (2017). Detecting Cancer Metastases on Gigapixel Pathology Images. *CoRR*.