Deep Residual Learning for Image Recognition

Subjects: Computer Science, Artificial Intelligence Contributor: Muhammad Shafiq , Zhaoguan Gu

In 2015, a deep residual network (ResNet) was proposed for image recognition. It is a type of convolutional neural network (CNN) where the input from the previous layer is added to the output of the current layer. Deep Residual Networks have recently been shown to significantly improve the performance of neural networks trained on ImageNet, with results beating all previous methods on this dataset by large margins in the image classification task.

deep residual learning for image recognition deep residual learning

image processing

image recognition

1. Introduction

Deep residual learning is a neural network architecture that was proposed in 2015 by He et al. ^[1] The paper *Deep* Residual Learning for Image Recognition has been cited many times and is one of the most influential papers in the field of computer vision. In this entry, researchers will survey the recent advances in deep residual learning. After discussing what deep residual networks are, researchers will review their properties, including stability and trainability. Next, researchers will discuss some recent applications of deep residual networks. Finally, researchers will provide their thoughts on future research directions in deep residual learning and end with open questions. This comprehensive survey looks at the current state of the art in deep learning for image recognition and proposes a new method, called deep residual learning, which offers significant improvements over existing methods. The author in ¹ provides a detailed overview of the proposed approach and its advantages. The proposed deep residual learning is computationally efficient as it has a low parameter number and uses simple backpropagation to reduce computation cost.

They also ^[2] suggest that there are applications other than just image recognition, such as translation and speech recognition, which could benefit from deep residual learning. Similarly, author ³ presents comparisons between different models with different architectures and finds that deep residual models always outperform other models. In addition, the author points out various challenges in applying the proposed deep residual learning. For instance, how do we deal with saturation and dropout? How do we deal with tasks like translation where there is less data available? The author concludes their paper by suggesting future research directions on how these challenges can be overcome. They point out that more details should be studied on combining deep residual learning with neural architecture search, spatial domain convolutions, constrained adversarial loss function and Gaussian-based generative models. Couso in 2018 in ^[2] proposes an alternative algorithm for maximizing likelihood rather than mean square error, which is currently used. They also suggest studying the proposed model for other computer vision tasks like face detection, segmentation, and object classification. They conclude their research study by pointing out limitations in proposed deep residual learning. They noted that compared to traditional methods, the proposed deep residual learning lacks computational efficiency when dealing with larger data sets and thus cannot scale up quickly enough. However, the authors point out that this problem can be solved by clustering input data into smaller subsets so only a subset of the total data needs to be processed during each iteration. Similarly, Feng et al. in ^[4] mention that deep residual learning, like any other unsupervised learning, requires a large amount of unlabeled data. They have done some experiments with a small amount of labeled data but they have not been able to get satisfactory results. The authors end their work by mentioning possible solutions: introducing few labels (which may require human intervention) or adding a fully supervised component. They also propose to create a dataset that contains images with predefined metadata and use the metadata as supervision.

The author in ^[5] ends their study by concluding that deep residual learning for image recognition is a promising direction in image recognition. They note that deep residual learning for image recognition is computationally efficient, more accurate and suitable for sparse data representations. They also emphasize the fact that deep residual learning for image recognition does not depend on complicated handcrafted features or the topographic organization of input data. However, the detailed workflow of researchers' proposed survey is shown in **Figure 1**.



Figure 1. Detailed Flowchart.

They conclude their paper by proposing directions for future research which include looking into combining deep residual learning with neural architecture search, spatial domain convolutions, constrained adversarial loss function and Gaussian-based generative models. In their study they also mention the need to find a solution to minimize the negative effect of data noise in deep residual learning. Mindy yang et al. in ^[6] proposes to create a dataset that contains images with predefined metadata and use the metadata as supervision. They show that deep residual learning for image recognition is computationally expensive because it is sensitive to high dimensional data and propose to combine deep residual learning with artificial intelligence techniques like reinforcement learning, and they discuss whether or not deep residual learning will help advance machine learning in general.

The study also notes that deep residual learning is successful because it can take advantage of large amounts of training data without requiring much hand engineering or task-specific feature engineering. The author of ^[7] proposes to develop a new framework which may be required when using deep residual learning. These tools are difficult to make assumptions about what they should do in terms of accuracy gains or performance losses due to complexity. The author states that deep residuals provide more accurate representations of object boundaries than traditional models and also allow for localization without global context. The basic structure is shown in **Figure 2**.



Figure 2. Basic Structure of DRN.

Zhu in ^[8] mentions how standard convolutional neural networks have limited the amount of parameters when compared to standard feedforward neural networks, which could explain why these types of models outperform them on certain tasks such as detection, localization, segmentation, tracking, and classification. The author mentions how difficult it would be to design different objective functions, with one per desired application. They propose an improved form of standard deep residual learning by combining batches into a single input batch and taking an average gradient across all images. The idea behind this is the hope that all datasets are similar enough so that averaging their gradients improves the gradient quality.

The result of implementing these changes was improved computational efficiency, while maintaining good accuracy levels which might make this model applicable to real world applications. The proposed improvements were shown to improve deep residual learning's ability to work well with large amounts of high dimensional data and thus make it useful for many applications. It would be interesting to see if these improvements in computing power can be applied to other fields besides image recognition. The author then goes on to talk about future steps and argues that deep residual learning's potential beyond image recognition needs to be explored.

2. What Is a Deep Residual Network?

In 2015, a deep residual network (ResNet) was proposed by the authors in ^[1] for image recognition. It is a type of convolutional neural network (CNN) where the input from the previous layer is added to the output of the current layer. This skip connection makes it easier for the network to learn and results in better performance. The ResNet architecture has been successful in a number of tasks, including image classification, object detection, and semantic segmentation. Additionally, since ResNets are made up of layers, these networks can be arbitrarily deep for an arbitrary level of spatial representation. There are various reasons for the success of the model: the large receptive fields that capture more information about each pixel in an image; the separation between the localization and classification stages; the computational efficiency at higher levels; the efficient encoding schemes with low-complexity arithmetic operations; and there is increased accuracy as features are extracted deeper into the network.

Despite these advantages, current ResNets are computationally very expensive. While modern GPUs can perform over one hundred million operations per second (Giga ops), a commonly used architecture of a fully connected layer with ten million weights takes more than two hours to train. This is why the authors in ^[9] propose to replace some fully connected layers by stochastic pooling layers and to reduce it from a 5×5 filter size to a 3×3 filter size.

In summary, deep residual learning for image recognition has been shown to be an effective method for image classification tasks. However, similar architectures have not yet been explored for other computer vision tasks such as semantic segmentation or object detection. There are several open problems that need further exploration when doing so, including computational efficiency at higher levels and training stability; adding skip connections; network depth versus complexity; biasing nonlinearities during training; input preprocessing issues such as batch normalization, data augmentation algorithms for improving accuracy for underrepresented classes such as at nighttime images versus daytime images using same classifier neural network through exploiting spatio-temporal

coherence; practicality of architectures; stability while small local minima not having significant impact upon generalization performance since big changes happen early vs. late in the sequence; this would allow for the concurrent tuning of different regions of parameters instead of completely independent ones.

The major issue with traditional CNNs is that they have to learn the entire feature map, which means that they need a huge number of parameters. This, in turn, means that they are very expensive to train and also slow to run.

ResNets are a family of neural networks that were proposed as an improvement over traditional CNNs. In particular, ResNets use skip connections (which I'll describe below), which allow them to be much smaller than traditional CNNs while still having similar performance. Skip connections can be used in any neural network architecture, but they are particularly useful for convolutional neural networks because they let you reuse parts of your feature map between layers in different positions.

Here researchers have a simple three-layer convolutional network with two convolutional layers followed by a pooling layer. The input is fed into layer 1, which performs its computation and outputs a feature map (which is just an array of numbers). Layer 2 then performs its own computation on top of layer 1's output and produces another feature map. This process repeats until it reaches the final layer.

Aryo Michael and Melki Garonga in 2021 ^[10] also proposed a new residual network with deep residual learning for image recognition, which integrates element-wise pooling with multi-scale features. Their approach combines depthwise separable convolution and deconvolution operations along with 2×2 and 3×3 convolutions to form different types of layers. For instance, layer 2 is made up of three layers: one that performs 4×4 convolution by padding its input images with zero; another that computes 2×2 deconvolution (i.e., transpose of spatial average); and a third layer that performs 3×3 convolution. In order to reduce computational costs, they replace some fully connected layers with more computationally efficient ones like stochastic pooling layers. The authors in proposed a new residual network with deep residual learning for image recognition. This is a hybrid model that incorporates both LSTM and CNNs. They show that their proposed architecture of outperforms.

In order to reduce computation costs, the authors in ^[11] propose replacing some fully connected layers with more computationally efficient ones like stochastic pooling layers. They propose a new residual network with deep residual learning for image recognition, which is a hybrid model that incorporates both LSTM and CNNs. This proposed architecture outperforms the ResNet-50 benchmark in terms of top-1 and top-5 error rates for the CIFAR10 dataset with comparable computational cost to the original ResNet-50. Furthermore, there are still a number of open problems that require more research, such as parallelizing for faster execution, using learned representations for transfer learning and sparse networks for reducing memory consumption, and using unsupervised feature extraction techniques to obtain meaningful high level descriptors and visual representations. Another idea is to use unlabeled and semi-labeled images for learning additional task-specific image descriptors and filters. Finally, deep residual learning for image recognition should be investigated for sequences of more than three frames in video and scene understanding.

Thus, the proposed architecture of deep residual learning for image recognition outperforms the ResNet-50 benchmark in terms of top-one and top-five error rates for the CIFAR10 dataset with comparable computational costs to the original ResNet-50. Future work should include investigations of how deep residual learning for image recognition might function with higher level applications such as scene understanding or video processing. This can potentially provide a strong foundation for leveraging deep residual learning for image recognition for future tasks such as human pose estimation. For the purposes of large-scale learning, deep residual learning for image recognition may be able to speed up supervised learning of heterogeneous datasets where the volumes of data and computation power available are limited. On a smaller scale, deep residual learning for image recognition may improve real-time responsiveness in autonomous vehicles and better respond to dynamic environments that include objects that appear, disappear, or change position.

There is potential with these models that is untapped due to many applications being based off of humans annotating their own photographs with different objects and labels. The development of detectors in machine systems will need human assistance to operate effectively. These detectors need to be trained effectively in order for them to take the place of human labor and assist industry ^[12] at greater levels. Developments have been proposed in methods for vision including deep residual learning for image recognition, where a model uses both neural networks and long short term memory recurrent neural networks (LSTMs). These proposed architectures are necessary because traditionally neural networks have had a very limited capability when it comes to sequential information. The capabilities for current machines are mostly found through continuous streams or fast frames, but most action happens over extended periods of time, which does not lend itself well for traditional algorithms used in vision models. These methods need improvement so that systems do not fall behind those humans who will continue to develop them beyond this point. Reserves need to be planned for the future, and this requires investments in training and developing computers that are equipped with sophisticated sensors. If a sufficient amount of investment is made in deep residual learning for image recognition, it could solve the problem of autonomous machines that cannot identify features in an environment. Human operators would then only input markers for certain features that would make machine systems more intelligent. Machine-learning models could then guickly identify objects and text from the map given to them by humans. Humans could maintain control until the system becomes proficient enough to support its users without supervision, but with a continuous stream of feedback from the AI system. This feedback loop for machine systems would help refine the AI's objectives and approach to specific problems, which in turn will allow for refined solutions for any objective. Machines can always become smarter than humans in some aspect of intelligence, but this shouldn't impede our ability to teach them what we know so they may learn faster than we ever could alone.

3. What Is Image Recognition?

Image recognition is a field of computer vision that deals with the identification and classification of objects in digital images. It is a subset of machine learning, which is a branch of artificial intelligence that deals with the design and development of algorithms that can learn from and make predictions on data. When you take a picture, it could be an image containing any number of items: dogs, cars, people, etc. There are countless other things in the world

besides these. The goal of image recognition is to assign categories to each one; so that when you upload an image to your social media feed or search Google Images you will get back information about what is in it and find out where else you might find them. For instance, if you were looking at this photo of someone holding their dog, researchers' application would recognize that there is a person in the photo and show their name as well as the name of their pet.

Image recognition is a problem within computer vision which refers to automatically detecting and understanding a wide range of objects in images. Computer vision can be seen as an artificial version of human sight or photography. There are several steps involved in image recognition. The first step is usually to convert an image into numbers that computers understand. An image contains hundreds of thousands (if not millions) of colors that are made up of red, green and blue (RGB) ^[13]. These colors are turned into data points which form vectors what researchers call images. A vector has three different values: one for each color channel. To reduce the size of the resulting vector and represent just a few shades of a particular color instead, methods use either linear classifiers or non-linear classifiers to create a predictive model that is able to classify new inputs.

Currently, most deep neural networks include residual connections that help propagate local gradients through many layers of nonlinear hidden units without exploding gradients. The future of deep residual learning is promising and worthy of further exploration. Deep residual learning is a form of reinforcement learning that builds on the success of residual networks. Deep residual networks have shown great potential in a variety of settings and have been used to solve image recognition tasks such as classification and semantic segmentation with high levels of accuracy.

The author in 2021 ^[14] concludes that deep residual learning offers significant improvements over traditional techniques such as VGG and ResNet on image recognition tasks, which they have explored with CNNs and RNNs. They note that future work should include investigating the effects of changing some parameters in the networks, exploring how deep residual networks adapt to sequential data, accounting for full system quality metrics, and furthering exploration into this topic.

In their paper, they detail multiple architectures including a novel global alignment- based network architecture combined with region proposal generation using two additional channels. They propose a global alignment approach between those two generators along with more structure to determine whether a given pixel was in motion or not. They determined that these global constraints made it easier for neural network methods to provide accurate proposals on image recognition tasks, allowing for higher accuracy and computational savings when compared against individual models trained without these constraints. The proposed methodology provides a high level of accuracy while being more computationally efficient than previous approaches. In the proposed frameworks, researchers present and analyze the main performance metrics for classification and semantic segmentation. The authors go on to describe their proposed methodologies in-depth, noting that all experimental datasets were gathered from public sources. Image recognition is a problem within computer vision, which refers to automatically detecting and understanding a wide range of objects in images.

4. Deep Residual Learning for Image Recognition

4.1. Deep Residual Learning Image Steganalysis

Image steganography is a technique that allows for the hiding of data in the image, and this data can be only visible when the image is modified. Deep residual learning image steganalysis is a technique that enables an attacker to find out what information has been moved in the image and how it has been moved. Deep residual learning image steganalysis is a new method to detect steganography. This method is based on deep residual networks that learn the local patch feature of images. The proposed security system is composed of three stages: pre-processing, feature extraction and classification.

Why is image steganalysis important? Image steganalysis is important because it can help protect against the unauthorized use of copyrighted material, help uncover inappropriate content hidden within images, and uncover potential security threats. What are some challenges in image steganalysis? The three most challenging aspects of image steganalysis are false positives, obtaining robust features from an image to differentiate between noise and data, and developing models that learn well from few training examples. There are several approaches to address these challenges including active learning approaches for extracting good features from an image; however, these approaches often involve expensive human labor or resources. The false positive rate has been shown to be high enough that many large-scale search engines don't even bother scanning for them because they're not worth the cost of storage space or computing power required to weed out all those extraneous pictures they come across while looking for matches. That's why the creators of TinEye, one such search engine, have suggested that machine learning researchers take this problem more seriously. The goal was to develop techniques to detect any personal information added covertly to an image without altering its appearance. For example, consider an image of someone standing on the beach holding up their child. By examining the pixels in an unaltered version of the picture, you might notice telltale signs that some numbers had been written over their head using pixelation techniques. Detecting these kinds of changes isn't easy to do by hand, so if you want a computer to do it you need a system capable of detecting very small changes made to individual pixels. Once ADRIAN detects tiny pieces of data embedded inside a photo, it will then highlight those areas with color differences around them and allow investigators to zoom in on precisely where there are visible changes. **Figure 3** shows the next sections.



Figure 3. Next Five Topics.

4.2. Deep Residual Learning Image Compression

Christian Rathgeb eta al in ^[15] studied the effect of image compression on the accuracy of deep learning models. They found that image compression can reduce the size of training datasets by up to 90% without any significant loss in accuracy. This is because deep learning models are able to learn the relevant features from data more effectively than shallower models. Image compression can also help speed up training times and reduce the amount of memory required to store training data. They looked at different ways of compressing images for use as inputs to convolutional neural networks (CNNs). They noted that arithmetic coding may provide an improvement over Huffman coding due to its ability to avoid rounding errors. However, it may be challenging for arithmetic coding-based methods to achieve similar results as Huffman coding-based methods when encoding high resolution images due to the large number of non-zero coefficients present within such images.

The authors in investigated whether human vision or machine vision algorithms could outperform one another when identifying objects in natural images. They performed their experiment using their own dataset of animal photographs they had manually annotated. They used LBP Features and COSINE scale space representations in order to compute similarities between pairs of images. Their experiments revealed that humans perform better than machine vision algorithms in tasks involving object recognition, localization, segmentation, etc., while machine vision performs better than humans in tasks involving detection and pose estimation. Machine vision algorithms are also less computationally expensive.

The authors in ^[16] studied how semantic segmentation can be employed to aid applications in various industries including medicine, self-driving cars, and surveillance. They concluded that computer vision models trained via unsupervised learning are capable of producing more accurate results than models trained via supervised learning. Moreover, it was discovered that training methods based on either small or large minibatches performed better than methods based on medium sized minibatches. Finally, it was found that unsupervised training techniques performed better than supervised ones when utilizing noisy labels for data labeling purposes.

This entry confirms that machine vision models based on unsupervised learning can perform just as well as those based on supervised learning, but have the added benefit of being quicker to train.

However, how semantic segmentation can be employed to aid applications in various industries including medicine, self-driving cars, and surveillance. They concluded that computer vision models trained via unsupervised learning are capable of producing more accurate results than models trained via supervised learning. Moreover, it was discovered that training methods based on either small or large minibatches performed better than methods based on medium sized minibatches.

4.3. Deep Residual Learning Image Restoration

Image restoration ^[17] is the problem of removing a uniform blur from an image. It is well-known that informationtheoretic approaches to this problem, based on the concept of a log-likelihood ratio operator, can be modelled well by deep neural networks (DNNs). Recent work has shown that DNNs trained on maximising generalisation performance can also be used to solve this task with remarkable effectiveness. This entry describes an extension of such an architecture; researchers call it Deep Residual Learning (DRL). DRL uses Riemannian geometry to minimise the cost function and achieve an optimal sparse approximation of the true posterior density.

The author in ^[18] studied the use of deep residual learning for image recognition. They found that it can effectively remove noise and improve the performance of image restoration models. In addition, they showed that deep residual learning can be used to improve the accuracy of image classification models. The authors also demonstrated that deep residual learning can be used to improve the performance of object detection models. Finally, they showed that deep residual learning can be used to improve the used to improve the accuracy of scene recognition models.

Similarly, the author in ^[19] showed that deep residual learning can be used to improve the accuracy of image retrieval models. Similarly, the author in ^[20] showed that deep residual learning can be used to improve the accuracy of video captioning models. The author in pointed out that deep residual learning presents an excellent alternative to traditional deep neural networks. Using deep-residual-learning may lead to improvements on accuracy or speed of inference at run time without compromising other objectives such as throughput or energy efficiency ^{[21][22]}. Similarly, they suggested that this method could be combined with variational methods to handle data sparsity and label noise. Similarly, deep residual learning presents an excellent alternative to traditional deep neural networks. Similarly, they also noted that using deep-residual-learning may lead to improvements in accuracy or speed of inference at run time without compromising other objectives such as throughput or energy neural networks. Similarly, they also noted that using deep-residual-learning may lead to improvements in accuracy or speed of inference at run time without compromising other objectives such as throughput or energy efficiency.

4.4. Deep Residual Learning Sensing Ct Reconstruction

Deep Residual Discriminator (Deep Residual Learning Sensing Ct Reconstruction) ^{[23][24]} is a deep learning model for helping radiologists classify and identify detections, which are placed in the CT image ^[25]. The goal of Deep Residual Discriminator is to improve radiologists' workflow by reducing the time it takes to scan and produce reports, while also improving detection efficacy.

The author in ^[26] studied the problem of image recognition using deep residual learning. They proposed a method that can be used to achieve high accuracy in image recognition tasks. The proposed method is based on the idea of using deep residual learning to improve the accuracy of image recognition models. The authors showed that their method can achieve better accuracy than the state-of-the-art methods on the ImageNet dataset. They also analyzed and evaluated various modifications to the deep residual network, and showed that it outperforms other approaches like dropout and batch normalization. The authors noted that there are some limitations to the study. Another limitation is that they only considered convolutional neural networks, which means that their conclusions may not extend to generative adversarial networks or recurrent neural networks. Nevertheless, their results indicate the promise of deep residual learning as an effective way to improve classification accuracy. For example, when comparing the traditional ReLU layer to ReLU + depthwise separable convolution layer for depth 16, the

accuracy improved ^[27]. When comparing a combination of all three layers (ReLU + depthwise separable convolution + kernel activation) versus ReLU alone at depths 16 and 32, then the combined layer performed significantly better. The results show that including deep residual layers within a neural network has significant benefits on performance. However, the use of many layers slows down computation speed. Future research should explore ways to train large deep networks faster while still achieving good accuracy.

In conclusion, deep residual learning was shown to provide promising improvement over standard architectures for image recognition. However, additional research needs to be done before researchers can make firm conclusions about how much improvement it offers and why this approach might work better than others currently available. A major limitation of this research is that only one type of neural network architecture was investigated, meaning that generalizations to other architectures need to be made cautiously. Also, little data were given regarding potential drawbacks to using deep residual learning for training deep networks such as increases in computational complexity or difficulty scaling up for very large networks. Finally, it would have been helpful if the researchers had examined whether specific optimization techniques were applied better with deep residual learning or without it; unfortunately, no such comparison was made here.

4.5. Deep Residual Learning Hyperparameter Optimization

Deep Residual Learning Hyperparameter Optimization ^[28] is a method for optimizing the hyperparameters of a Deep Residual Learning model. When optimizing the parameters of a deep residual network for image recognition, there are many factors to consider. The first is the depth of the network. Deeper networks have more layers and can therefore learn more complex features. Shallower networks, on the other hand, are faster to train and may be more efficient in terms of memory usage. Another important parameter is the width of the network, which refers to the number of neurons in each layer. Wider networks can learn more complex features, but are also more expensive to train. Finally, the learning rate is an important parameter that controls how quickly the network learns from data. A higher learning rate means that the network learns faster, but may also be more likely to overfit on the training data. If researchers use a standard gradient descent algorithm to optimize the parameters of a neural network such as this one, it is not guaranteed that researchers will find good local minima. It would be better if researchers had some way of knowing what global minima might look like before starting optimization so that researchers could search intelligently instead of randomly guessing where they might be.

Recent work has attempted to do this by using linear classifiers such as SVM's or k-nearest neighbors to label different regions in feature space based on whether they corresponded with positive or negative examples respectively. Then, these labeled regions were used as starting points for gradient descent optimization, providing researchers with potential solutions near these regions instead of random ones. While this technique provides researchers with valuable insight into where to start optimization, it still suffers from problems because these labels aren't perfect. For example, the labels are only correct 50% of the time, meaning that researchers' search process is less accurate than researchers would hope. More research is needed to improve upon this technique and make sure that it gives reliable results every time. However, despite its imperfections, this approach does provide new insights into the problem of hyperparameter optimization for deep networks and may help lead to improve

methods in the future. One of the main problems for optimization algorithms is that the gradients become too small when you go down into lower layers. Therefore, all traditional optimizers try to converge at a single minimum when in reality there are multiple local minima. One solution proposed was using non-convex strategies like Bayesian Optimization (BO) and genetic algorithms (GA). In BO you estimate gradients at many different points simultaneously, while GA does not rely on gradient information at all. Since both techniques show promise, further research should explore their performance in practice.

There are many factors to consider when optimizing the parameters of a deep residual network for image recognition. This entry discussed three key considerations: the depth of the network, width of the network, and learning rate. There is currently much ongoing research exploring novel ways to automate this process; however, no clear winner has emerged yet. There are many issues that need to be addressed, including gradient size and convexity. Other approaches include Bayesian optimization and genetic algorithms.

4.6. Very Deep Convolutional Network Image Recognition

In the past few years, convolutional neural networks (CNNs) have revolutionized image recognition by achieving unparalleled accuracy on benchmark datasets. A key ingredient in this success has been the use of very deep CNNs ^{[29][30]}, which are able to learn rich representations of images.

Recently, the author in ^[31] proposed a deep recursive residual network (DRRN) to address the problem of image super-resolution. The proposed DRRN consists of three stages: (1) the downsizing stage, (2) the upsampling stage, and (3) the reconstruction stage. In the downsizing stage, DRRN uses a convolutional neural network (CNN) to downscale an input image into a fixed size using a single channel. Then, in the upsampling stage, DRRN applies an upsampling layer to generate several intermediate images from the downsampled image. Finally, in the reconstruction stage, DRRN uses two CNNs to reconstruct an output image from these intermediate images. However, it is hard to find a suitable dataset for super-resolution. In most cases, the input images are different from each other. Therefore, researchers cannot use the same CNN architecture across different resolutions. Researchers need to design an architecture that can accommodate different resolutions without overfitting to lowresolution images (which is hard). You might want to try using a ResNet-like architecture with multiple residual blocks (ResNet has multiple branches), which may help you achieve good performance. The problem with transfer learning is that it works well in the sense that it is often able to learn a useful representation from a large amount of data, but it does not necessarily learn the best representation for the task at hand. For example, if you are trying to use transfer learning for face recognition and you have trained on a large number of faces, then you might find that your transfer learned model does not perform as well as one that was trained directly on faces. Similarly, the author in [32] proposes a new method for matching software-generated sketches with face photographs using a very deep convolutional neural network (CNN). This method uses two different types of networks: one network is trained on face photographs and another network is trained on sketches. The two networks are combined into one network by using transfer learning. Their experiments show that their method outperforms other state-of-the-art methods in terms of accuracy and generalization capability.

Shun Moriya and Chihiro Shibata ^[33] propose a novel transfer learning method for very deep CNNs for text classification. Researchers' main contribution is a new evaluation method that compares the proposed transfer learning method with two existing methods, namely fine-tuning and feature handover. They also propose a new model ensemble approach to improve the performance of researchers' models by using the best performing model from each ensemble member as an additional feature. Researchers' experiments on five public datasets show that researchers' approach outperforms previous methods and gives competitive results when compared with other state-of-the-art methods.

Similarly, Afzal et al. in ^[34] present the first investigation of how very deep Convolutional Neural Networks (CNN) can be used to improve document image classification. They also study how advanced training strategies such as multi-network training and model compression techniques can be combined with very deep CNNs to further improve performance. Their results show that very deep CNNs are able to outperform shallow networks, even when using a relatively small amount of training data. They also find that multi-network training significantly improves performance over single-network training, especially for very deep CNNs. Finally, they demonstrate that model compression techniques such as quantization and binarization can be combined with very deep CNNs to achieve an additional 5% reduction in error rate with only a small loss in accuracy. However, they should highlight that their model achieves state-of-the-art performance in document image classification. However, they don't provide any quantitative results to support this claim. They could easily add some performance metrics (e.g., F1 score) on top of their results, and this would make the paper more convincing. The next five sections are shown in **Figure 4**.



Figure 4. Next Five Topics.

4.7. Deep Residual Networks Accelerator on FPGA

A recent survey by ImageNet found that deep residual networks (ResNets) have become the state-of-the-art in image recognition ^[35]. However, training and deploying these models can be prohibitively expensive. FPGAs ^[36] offer a high degree of parallelism and energy efficiency, making them an attractive platform for accelerating deep neural networks. In this entry, researchers will survey the literature on FPGA-based acceleration of deep residual

networks. Researchers will discuss the challenges involved in training and deploying these models on FPGAs, and researchers will survey the current state-of-the-art in FPGA-based deep learning accelerators. Worthy of note are the paper from Hui Liao et al., which presents a systematic comparison between CPU-based and GPU-based training of ResNets.

4.8. Resnet Models Image Recognition

In 2015, a new deep learning model known as a deep residual network (ResNet) was introduced by researchers at Microsoft ^[1]. This model has guickly become the state-of-the-art for image recognition tasks. These networks are now part of Convolutional Neural Networks (CNNs) [37]. They have been used to achieve world records in object classification and detection in many large-scale competitions such as ImageNet, ILSVRC, COCO and PASCAL VOC. Furthermore, they also achieved competitive results on various 3D shape estimation problems. These models are computationally expensive because they require billions of parameters and need hundreds of millions of training images. Consequently, this has led to some speculation that they will never be used outside academia or research laboratories due to their high computation cost; however, recent developments like adopting modern graphics processing units (GPUs) may bring down these costs significantly in the near future. The first significant attempt to train such models using GPUs came from NVIDIA's [38] deep learning framework CUDA [39] back in 2014. However, Nvidia soon found out that there were limitations of scaling the GPU implementations to larger Resnet architectures [34][40], and it was difficult to provide a stable environment for training. The researchers then turned towards software written specifically for use on GPUs which could take advantage of newer hardware capabilities. It is anticipated that these new frameworks will offer significant improvements over their predecessors because they provide an interface between multi-core CPUs and GPUs while also offering data preprocessing functions which are necessary when working with large datasets. One such framework is called Intel Integrated Performance Primitives Library (Intel IPP), which offers a variety of different functions including matrix multiplication, convolutions, etc. Thus far it has shown good performance on small and medium sized datasets but not so much on larger ones. Furthermore, another promising library that can also take advantage of Intel processors' Single Instruction Multiple Data (SIMD) technology [41] is Eigen and its variants like HKLMSVD or GEMM. This library contains algorithms designed for Numerical Linear Algebra that would work well with ResNets. These implementations have shown excellent performance on both ImageNet and Cityscapes benchmarks, yet they still remain challenging to parallelize without sacrificing too much accuracy.

The time spent training image recognition models has fallen dramatically since 2010, owing to increasing computational power and the availability of massive labeled datasets. In 2016, Google trained a model within six days using 8 TPUs (Tensor Processing Units). Another breakthrough was achieved by Baidu's Sunway TaihuLight ^[42] computer, which is based on China's national design blueprint created in 2013 ^[43]. The three grand challenges in fundamental science are high performance computing, brain science, and quantum computing.

4.9. Shrinkage Network Image Recognition

Shrinkage Network Image Recognition ^{[44][45]} is an important tool for image recognition. This method uses a deep residual learning framework to achieve state-of-the-art performance on various image recognition tasks. The authors mention that the networks are composed of three key components, namely depthwise convolution, max pooling and subsampling layers.

In particular, the depthwise convolution [46] performs feature extraction by mapping input images onto filter responses at different depths. Max pooling captures spatial information by aggregating features over a fixed window size across the input channels in both spatial dimensions and selecting top k features from each window. Finally, subsampling layers are responsible for reducing network size while maintaining its accuracy via training the network on reduced resolution images (or upsampled or downsampled images). They mention that using this architecture can reduce training time from hours to minutes per epoch without affecting accuracy significantly (or even improving it). Furthermore, they mention that there have been some recent improvements with regard to previous versions such as the use of LSTM units instead of RNNs [47][48], which can improve robustness against adversarial attacks. One other interesting point mentioned was how the researchers used RGBD data to better understand how humans perceive color and objects. They found that humans tend to see colors mostly in the midspectrum where red, green and blue meet, so they created special networks that mimic human perception of color when training them. Another discovery is that researchers find objects more easily if they are located near edges rather than in cluttered areas. To take advantage of this finding, their model predicts two probabilities, one corresponding to the probability of detecting an object in the cluttered region and another corresponding to the probability of detecting an object at the edge regions. The final model achieved good results on classification tasks like labeling dogs versus cats and types of food (bananas vs. apples).

On small datasets like CIFAR-10 ^[49], a new approach called Instance Normalization achieves competitive results, but when applied to larger datasets like ImageNet large gains were obtained.

4.10. Tomography Images Deep Learning and Transfer

Hao et al. in ^[50] studied the application of deep learning to tomography images and found that the proposed method can effectively improve the recognition performance. Similarly, deep residual learning has also been applied to image recognition tasks with promising results. In this entry, researchers review the recent progress made in deep residual learning for image recognition. Researchers first introduce the general framework of deep residual learning and then discuss its application to various image recognition tasks. Finally, researchers summarize the challenges and future directions of deep residual learning for image recognition. In summary, deep residual learning has shown a promising application to image recognition tasks with relatively strong results.

Deep residual networks are composed of two parts: (1) dense layers and (2) downsampling layers ^[51], which aim at restoring lost details by averaging information from neighboring locations or different depths in the same layer. Densely connected layers form an intermediate representation which is stored for later use. Downsampling layers help produce more concise representations which are easier to train. There have been multiple successes applying deep residual networks in image recognition, as shown below. One example is image segmentation. A pre-trained

deep residual network was used to build a boundary map which was applied onto input data. The boundary map allows for the accurate labeling of different objects in the scene while preserving intricate features like edges and contours.

Similarly, another study used unsupervised pretraining [52][53] followed by supervised fine-tuning to classify six vehicle classes (categories). The classification accuracy obtained using the resulting deep network is close to 99%. It should be noted that these experiments have not yet gone beyond 10 training epochs, showing there may be room for improvement. Despite this, deep residual learning has already demonstrated its promise in providing increased accuracy over previous methods. For example, recent research shows that a combined approach of fully convolutional neural networks and 3D convolutional networks significantly outperforms both other approaches when classifying knee joint status from MRI scans. Similarly, it was recently shown that adding depth information leads to improved classification accuracies for facial expression detection, with smile detection achieving 96% accuracy and sulk detection achieving 88% accuracy when compared to 68% and 52% respectively without depth information. Moreover, even higher accuracies were obtained using regularized discriminative models such as ensemble perceptrons or boosting discriminative neural networks. Similarly, combining variational autoencoders with a generative adversarial network resulted in significant improvements for reconstructing speech from laryngeal articulations. Lastly, it should be noted that the field of computer vision has benefited greatly from the development of deep learning algorithms. Indeed, deep residual networks are one of many novel techniques that have been developed and tested over the past few years. Future studies will need to consider how best to leverage all aspects of their model architecture (such as their number of hidden layers), how much data they require during training, whether they require supervision or not during training and what kind they require during testing (for instance label noise). Furthermore, new tasks are needed to test the limitations of deep residual learning. As mentioned earlier, deep residual networks have been successfully applied to image recognition tasks. However, little work has been done in applying deep residual networks to natural language processing or object detection tasks. Similarly, further work is needed in understanding how transferable the learned weights are from task to task and if they require some sort of reconstruction process after being applied to a new task.

Deep residual learning is a powerful tool for applications requiring high levels of accuracy as well as robustness against changes in the distribution of inputs during training or testing. Deep residual networks represent a compelling alternative for dealing with visual recognition problems where datasets are limited or costly to collect.

4.11. Hybrid Deep Learning

Hybrid Deep Learning ^{[54][55]} is a proven method used to build the models. It can overcome the limitations of both traditional deep learning and reinforcement learning. This technique has been applied in complex real-world problems with impressive results. Hybrid Deep Learning can be used for any kind of signal processing task and it is going to be more important for new emerging applications such as self-driving vehicles or robotics.

In recent years, hybrid deep learning architectures ^[56] have been proposed to take advantage of the strengths of both CNNs and RNNs. The most successful hybrid models are based on a deep residual learning (ResNet)

framework proposed by He et al. ^[1]. Similarly, Deep Residual Learning was studied in Deep Residual Learning and ^[1] explored how ResNets can improve the performance of traditional supervised neural networks when combined with automatic feature engineering they investigated how ResNets can improve the performance of traditional supervised neural networks when combined with automatic feature engineering Author in found that they could learn low level features like contours, edges, corners and blobs just as well if not better than humans. The authors of ^[57] also explored practical challenges associated with applying deep resnets to computer vision problems such as working around computational limitations. In their study they also explored practical challenges associated with applying deep resnets to computer vision problems such as working around computational limitations. The authors concluded by discussing future directions, where they will investigate reinforcement learning on top of residual nets. They noted the promise of combining deep memory networks with other machine learning techniques such as dropout, autoencoders and generative adversarial networks to produce a powerful new generation of models. The details of the next chapter are shown in **Figure 5**.



Figure 5. Next Five Topics.

4.12. Deep Learning Architectures

In recent years, deep learning has made tremendous progress in the field of image recognition. The main contribution of this entry is a comprehensive survey of deep residual learning (ResNets), which is a state-of-the-art deep learning architecture for image recognition ^[58]. The author in ^[2] studied the effect of different depths of ResNets on the classification accuracy ^{[59][60]}. They found that deeper networks are more accurate than shallower networks. However, they also found that there is a diminishing return in accuracy as the network depth increases. Besides deepening ResNets, researchers have also investigated adding residual blocks to shallow models. The idea behind adding these blocks is to take advantage of the invariance properties of convolutional neural networks by using small images to compensate for large changes in input size and shape. These blocks help to prevent overfitting and underfitting problems in models with few parameters and a large number of filters. A very simple example includes replacing just one or two layers at the end of a shallow model with their corresponding ones from a deep model. One such extension is called SqueezeNet ^[61], which replaces only one layer with its corresponding layer from an entire deep CNN inception module. A further extension comes from Google Brain called

CheckerboardNet ^[62], which replaces all layers with their counterparts from an entire deep CNN inception module. Another type of extension comes from Microsoft Research, which consists of a series of hybrid architectures: Stacked Connected Convolutional Networks, Transformer Networks, and Self-Attention Models. All three architectures propose combining residual units with feature extraction units. Stacked Connected Convolutional Networks stack two types of networks on top of each other: (1) a set of regular deep networks where each level performs feature extraction; and (2) a single dense network containing up to 10 times fewer parameters than the previous level but still extracting features within each level. Similarly, Transformer Networks stack two types of models: (1) full-connected recurrent encoder-decoder pairs and (2) transformer blocks that replace groups connections between nodes while performing feature extraction. Lastly, Self-Attention Models stack two types of models: (1) Autoregressive Encoder Decoder Models; and (2) self-attention modules that extract features. Experiments conducted show that the proposed architectures produce better results than those without stacking residuals with extra modules. There are, however, various drawbacks associated with these architectures. Firstly, the training process can be time consuming. Secondly, there can be some significant increase in computation requirements since the same operations need to be performed repeatedly across different network layers. Finally, most of these proposals use computationally expensive attention mechanisms which increase both memory and computational requirements. For example, memory usage for Stacked Connected Convolutional Networks goes up from mm^2 to mm^3 . Self-attention models' computational cost per batch prediction goes up from O(1) to O(2). Attention mechanisms can provide robustness against adversarial examples, but it seems that it would be necessary to combine them with additional regularization techniques.

4.13. Deep Learning System

In recent years, deep learning systems ^[63] have achieved great success in many fields, including image recognition. One of the most successful deep learning models is the deep residual network (ResNet), which was proposed in 2015 ^[1]. Since then, ResNets have been widely used in various image recognition tasks and have shown state-of-the-art performance. However, there are still some limitations that need to be addressed: they are not computationally efficient and they can be easily fooled by adversarial examples. Recently, researchers have started to explore new methods to improve these shortcomings. Some promising approaches include changing the activation function from ReLU to ELU or SELU; using a memory module before each layer; using data augmentation techniques during training; and pre-training with a large dataset followed by fine tuning with a small dataset. The paper also reviews other methods such as attention-based networks, Generative Adversarial Networks, Convolutional Neural Networks, GANs, etc. The authors suggest future work to explore the aforementioned topics and make recommendations on how best to use deep residual networks for specific tasks. They also propose research directions in developing more powerful architectures based on deep residual networks, which may solve the problems faced by existing methods. Furthermore, in order to better utilise all the available resources, it is necessary to create an open-source software library which supports parallel computing. Additionally, novel datasets should be developed because current datasets only cover a limited number of object categories. When creating these datasets, researchers should take into account not just the pixel information but also the metadata. Furthermore, there needs to be a way to analyse data automatically without human involvement so that it is scalable and accurate. Lastly, future work should focus on studying what type of neural architecture would suit specific tasks such as semantic segmentation and pose estimation.

The main limitation of deep residual networks is that they are computationally inefficient and can be easily fooled by adversarial examples. There have been attempts to address these issues such as using a memory module before each layer, using data augmentation techniques during training, and pre-training with a large dataset followed by fine tuning with a small dataset. As for the future of deep residual networks for image recognition, it is hard to predict where they will go next. There have been studies on trying to combine deep residual networks with generative adversarial networks which yield encouraging results and hope for further developments in this area. A concern with these networks is that they are mainly trained on supervised tasks and it is not yet clear if they can be trained for reinforcement learning. Research should also look at the connections between deep residual networks and recurrent networks with long short-term memory units, for instance, in constructing deeper layers. Regarding the issue of efficiency, there have been significant advances such as combining ResNets with multi-level attention networks to increase the computational efficiency. The result is a huge reduction in required parameters and thus computation time. It remains to be seen how well these networks perform in comparison to standard models.

4.14. Deep Residual Learning Persistent Homology Analysis

Despite its great success, deep learning has been criticized [64] for being a black box. In this entry, researchers take a step towards understanding the inner workings of deep neural networks by performing a persistent homology analysis of the feature representations learned by a state-of-the-art deep residual network. These findings provide new insights into how convolutional and fully connected layers learn to extract increasingly abstract features from raw input data. They also have important implications for training deep networks because they suggest that in order to maximize performance, researchers should train them so as to push them deeper. Moreover, researchers' analysis gives rise to three new models for fast-forwarding through a sequence of images at different depths: (1) forward prediction: computing a feature representation at one depth predicting the next depth; (2) backward propagation: computing the representation at one depth propagating backwards up through previous layers; and (3) hybrid forward/backward propagation: running both forward and backward propagation steps simultaneously but treating each layer independently. Researchers' experimental results demonstrate that these models can significantly outperform traditional backpropagation when applied to videos, suggesting that future work may focus on developing efficient methods for training video processing tasks such as image segmentation and caption generation. The final result in this entry will be a continuation of the discussion found in Paper 2: A comprehensive survey on deep residual learning for image recognition concludes with two models for fast-forwarding through a sequence of images at different depths:

- Forward Prediction: computing a feature representation at one depth and predicting the next depth
- Backward Propagation: computing the representation at one depth propagating backwards up through previous layers.

- Hybrid Forward/Backward Propagation: Running both forward and backward propagation steps simultaneously but treating each layer independently.

In conclusion, despite all of the progress made thus far, there are still many challenges ahead for deep networks to overcome before they become commonplace in society. Researchers expect that there will be continued innovation and improvements over time leading to better computers, better algorithms, and ultimately better models for improving human life.

Researchers' hope is that these survey papers provide an overview from which future researchers can gain insight into approaches to tackling issues in computer vision and other areas of machine learning. Furthermore, researchers' goal is not simply to describe new concepts but also to hopefully inspire a new generation of researchers who wish to continue on with their own research endeavors. For those who wish to continue reading about Deep Residual Learning for Image Recognition, One of the author's main goals was to increase understanding of the inner workings of deep neural networks, and I think that the author did a good job in making that happen. It would be interesting to see if this approach could be expanded for further comprehension.

4.15. Deep Residual Learning Pest Identification

Deep residual learning is a neural network that is used to learn image representations. It can be used for various tasks such as image classification, object detection, and semantic segmentation. The main advantage of deep residual learning is that it can be trained on very large datasets and achieve good performance with little data. Similarly, Deep Residual Learning for Image Recognition authors studied the effects of applying deep residual learning on pest identification [65]. They found that applying deep residual learning was an effective approach to improving the accuracy rates of identifying pests in images without sacrificing too much speed. Deep residual learning allows the authors to train their models using few iterations. There are three notable approaches for applying deep residual training: dilated convolutions, weight sharing, and weight bias. Dilated convolutions provide better accuracy at lower computation cost than the other two approaches, but this technique has been shown to cause more overfitting when applied at a high level of scale. Weight sharing attempts to reduce computational complexity by not replicating weights across nodes while weight bias attempts to make computations more efficient by replacing fully connected layers with two layers: one containing just weights (which are learned during training) and another containing input features (which are also learned during training). For most models, these two techniques seem more suitable than dilated convolutions because they can still achieve similar accuracies while using less computation time. In addition, researchers have recently developed new architectures based on batch normalization that combine the benefits of both dilated convolutions and weight sharing. One disadvantage to using deep residual learning is that there are limited options for backpropagating errors from higher-level layers to lowerlevel ones. Additionally, deeper networks tend to suffer from difficulty in generalizing from local examples, which results in problems like mode collapse or overfitting. One way around this problem is to use batch normalization. Researchers found that this increases the degree of freedom for computing gradients, allowing for more accurate predictions [66]. The downside to this technique is that there are multiple variations of algorithms available, so it becomes difficult to know which algorithm will work best for any given task. A survey about deep residual learning

showed that there were challenges in trying to implement the technique when real-time inference was required, because certain parts of standard CPUs could not process the large number of multiplications quickly enough. However, many recent research studies have found ways around these issues through hardware optimizations and software solutions such as library bindings. Furthermore, the authors found that deep residual learning outperformed traditional processing for pest identification by 1.4%. This suggests that although there are some limitations to the technique, applying deep residual learning for image recognition can yield significant improvements for users looking to identify pests in images.

Deep residual learning can be implemented for image recognition in multiple different ways. Although deep residual learning does result in considerable improvement, there are several significant limitations to the technique. Deep residual learning works best when there is high resolution, large amounts of labeled datasets, and modern hardware that is optimized for implementing deep learning. Finally, it may take a while to optimize the dataset depending on how new the technology is. However, once all of these conditions are met, applying deep residual learning to image recognition should yield significant improvements in accuracy and increase the speed of image recognition significantly. Nevertheless, more testing is needed in order to gain a clearer picture on whether or not deep residual learning is suited for all applications that require analyzing and classifying images. Further research should be done to determine if applying deep residual learning to other application domains yields similar benefits as seen with image recognition.

References

- He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
- 2. Couso, I.; Dubois, D. A general framework for maximizing likelihood under incomplete data. Int. J. Approx. Reason. 2018, 93, 238–260.
- 3. Liang, Y.; Peng, W.; Zheng, Z.-J.; Silvén, O.; Zhao, G. A hybrid quantum–classical neural network with deep residual learning. Neural Netw. 2021, 143, 133–147.
- Feng, Z.; Nie, D.; Wang, L.; Shen, D. Semi-supervised learning for pelvic MR image segmentation based on multi-task residual fully convolutional networks. In Proceedings of the 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018; pp. 885–888.
- 5. Li, L.; Jin, W.; Huang, Y. Few-shot contrastive learning for image classification and its application to insulator identification. Appl. Intell. 2021, 52, 6148–6163.
- 6. Yang, M.; Thung, G. Classification of Trash for Recyclability Status. CS229Project Rep. 2016, 2016, 3.

- Karar, M.E.; Hemdan, E.E.-D.; Shouman, M.A. Cascaded deep learning classifiers for computeraided diagnosis of COVID-19 and pneumonia diseases in X-ray scans. Complex Intell. Syst. 2020, 7, 235–247.
- 8. Zhu, J.; Chen, H.; Ye, W. A Hybrid CNN–LSTM Network for the Classification of Human Activities Based on Micro-Doppler Radar. IEEE Access 2020, 8, 24713–24720.
- 9. FPGA Acceleration of Convolutional Neural Networks; Nallatech: Camarillo, CA, USA, 2017.
- 10. Michael, A.; Garonga, M. Classification model of 'Toraja' arabica coffee fruit ripeness levels using convolution neural network approach. ILKOM J. Ilm. 2021, 13, 226–234.
- Al-Kharraz, M.S.; Elrefaei, L.A.; Fadel, M.A. Automated System for Chromosome Karyotyping to Recognize the Most Common Numerical Abnormalities Using Deep Learning. IEEE Access 2020, 8, 157727–157747.
- 12. Avtar, R.; Tripathi, S.; Aggarwal, A.K.; Kumar, P. Population–Urbanization–Energy Nexus: A Review. Resources 2019, 8, 136.
- 13. Brachmann, E.; Rother, C. Visual Camera Re-Localization from RGB and RGB-D Images Using DSAC. IEEE Trans. Pattern Anal. Mach. Intell. 2021, 44, 5847–5865.
- 14. Akhand, M.; Roy, S.; Siddique, N.; Kamal, A.S.; Shimamura, T. Facial Emotion Recognition Using Transfer Learning in the Deep CNN. Electronics 2021, 10, 1036.
- 15. Rathgeb, C.; Bernardo, K.; Haryanto, N.E.; Busch, C. Effects of image compression on face image manipulation detection: A case study on facial retouching. IET Biom. 2021, 10, 342–355.
- Siam, M.; Elkerdawy, S.; Jagersand, M.; Yogamani, S. Deep semantic segmentation for automated driving: Taxonomy, roadmap and challenges. In Proceedings of the 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 16–19 October 2017.
- 17. Zhang, K.; Li, Y.; Zuo, W.; Zhang, L.; Van Gool, L.; Timofte, R. Plug-and-Play Image Restoration with Deep Denoiser Prior. IEEE Trans. Pattern Anal. Mach. Intell. 2021; early access.
- 18. Sangeetha, V.; Prasad, K.J.R. Deep Residual Learning for Image Recognition Kaiming. Indian J. Chem.-Sect. B Org. Med. Chem. 2006.
- 19. Cheng, S.; Wang, L.; Du, A. An Adaptive and Asymmetric Residual Hash for Fast Image Retrieval. IEEE Access 2019, 7, 78942–78953.
- Fujii, T.; Sei, Y.; Tahara, Y.; Orihara, R.; Ohsuga, A. "Never fry carrots without chopping" Generating Cooking Recipes from Cooking Videos Using Deep Learning Considering Previous Process. Int. J. Netw. Distrib. Comput. 2019, 7, 107.

- Avtar, R.; Sahu, N.; Aggarwal, A.K.; Chakraborty, S.; Kharrazi, A.; Yunus, A.P.; Dou, J.; Kurniawan, T.A. Exploring Renewable Energy Resources Using Remote Sensing and GIS—A Review. Resources 2019, 8, 149.
- Avtar, R.; Komolafe, A.A.; Kouser, A.; Singh, D.; Yunus, A.P.; Dou, J.; Kumar, P.; Das Gupta, R.; Johnson, B.A.; Minh, H.V.T.; et al. Assessing sustainable development prospects through remote sensing: A review. Remote Sens. Appl. Soc. Environ. 2020, 20, 100402.
- 23. Fu, Z.; Tseng, H.W.; Vedantham, S.; Karellas, A.; Bilgin, A. A residual dense network assisted sparse view reconstruction for breast computed tomography. Sci. Rep. 2020, 10, 21111.
- Wu, W.; Hu, D.; Niu, C.; Broeke, L.V.; Butler, A.P.; Cao, P.; Atlas, J.; Chernoglazov, A.; Vardhanabhuti, V.; Wang, G. Deep learning based spectral CT imaging. Neural Netw. 2021, 144, 342–358.
- 25. Jalali, Y.; Fateh, M.; Rezvani, M.; Abolghasemi, V.; Anisi, M.H. ResBCDU-Net: A Deep Learning Framework for Lung CT Image Segmentation. Sensors 2021, 21, 268.
- 26. Chalasani, P. Lung CT Image Recognition using Deep Learning Techniques to Detect Lung Cancer. Int. J. Emerg. Trends Eng. Res. 2020, 8, 3575–3579.
- 27. Cui, B.; Dong, X.-M.; Zhan, Q.; Peng, J.; Sun, W. LiteDepthwiseNet: A Lightweight Network for Hyperspectral Image Classification. IEEE Trans. Geosci. Remote Sens. 2021, 60, 1–15.
- Jafar, A.; Myungho, L. Hyperparameter Optimization for Deep Residual Learning in Image Classification. In Proceedings of the 2020 IEEE International Conference on Autonomic Computing and Self-Organizing Systems Companion (ACSOS-C), Washington, DC, USA, 17–21 August 2020.
- 29. Qian, Y.; Bi, M.; Tan, T.; Yu, K. Very Deep Convolutional Neural Networks for Noise Robust Speech Recognition. IEEE/ACM Trans. Audio Speech Lang. Process. 2016, 24, 2263–2276.
- 30. Wang, R.; Tao, D. Training Very Deep CNNs for General Non-Blind Deconvolution. IEEE Trans. Image Process. 2018, 27, 2897–2910.
- Tai, Y.; Yang, J.; Liu, X. Image super-resolution via deep recursive residual network. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017.
- Galea, C.; Farrugia, R.A. Matching Software-Generated Sketches to Face Photographs With a Very Deep CNN, Morphed Faces, and Transfer Learning. IEEE Trans. Inf. Forensics Secur. 2017, 13, 1421–1431.
- Moriya, S.; Shibata, C. Transfer Learning Method for Very Deep CNN for Text Classification and Methods for its Evaluation. In Proceedings of the 2018 IEEE 42nd Annual Computer Software and Applications Conference (COMPSAC), Tokyo, Japan, 23–27 July 2018; Volume 2.

- 34. Afzal, M.Z.; Kolsch, A.; Ahmed, S.; Liwicki, M. Cutting the Error by Half: Investigation of Very Deep CNN and Advanced Training Strategies for Document Image Classification. In Proceedings of the 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), Kyoto, Japan, 9–15 November 2017.
- 35. Bashir, S.M.A.; Wang, Y.; Khan, M.; Niu, Y. A comprehensive review of deep learning-based single image super-resolution. PeerJ Comput. Sci. 2021, 7, e621.
- 36. Bao, C.; Xie, T.; Feng, W.; Chang, L.; Yu, C. A Power-Efficient Optimizing Framework FPGA Accelerator Based on Winograd for YOLO. IEEE Access 2020, 8, 94307–94317.
- Lim, H.K.; Kim, J.B.; Heo, J.S.; Kim, K.; Hong, Y.G.; Han, Y.H. Packet-based network traffic classification using deep learning. In Proceedings of the 2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC), Okinawa, Japan, 11–13 February 2019.
- 38. Available online: https://cyberleninka.ru/article/n/reshenie-zadach-vychislitelnoy-gidrodinamiki-sprimeneniem-tehnologii-nvidia-cuda-articlehead-tehnologiya-nvidia-cuda-v-zadachah/viewer (accessed on 2 September 2022).
- 39. NVIDIA. Cuda C Best Practices Guide; Nvidia Corp.: Santa Clara, CA, USA, 2015.
- 40. Yasin, S.; Iqbal, N.; Ali, T.; Draz, U.; Alqahtani, A.; Irfan, M.; Rehman, A.; Glowacz, A.; Alqhtani, S.; Proniewska, K.; et al. Severity Grading and Early Retinopathy Lesion Detection through Hybrid Inception-ResNet Architecture. Sensors 2021, 21, 6933.
- Li, Y.; Xie, P.; Chen, X.; Liu, J.; Yang, B.; Li, S.; Gong, C.; Gan, X.; Xu, H. VBSF: A new storage format for SIMD sparse matrix–vector multiplication on modern processors. J. Supercomput. 2019, 76, 2063–2081.
- 42. Li, R.; Wu, B.; Ying, M.; Sun, X.; Yang, G. Quantum Supremacy Circuit Simulation on Sunway TaihuLight. IEEE Trans. Parallel Distrib. Syst. 2019, 31, 805–816.
- 43. Guarnieri, M. Trailblazers in Electromechanical Computing . IEEE Ind. Electron. Mag. 2017, 11, 58–62.
- Li, Y.; Chen, H. Image recognition based on deep residual shrinkage Network. In Proceedings of the 2021 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA), Guangzhou, China, 14–16 May 2021.
- 45. Yang, Z.; Wu, B.; Wang, Z.; Li, Y.; Feng, H. Image Recognition Based on an Improved Deep Residual Shrinkage Network. SSRN Electron. J. 2022; in press.
- 46. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.

- 47. Javed, A.R.; Usman, M.; Rehman, S.U.; Khan, M.U.; Haghighi, M.S. Anomaly Detection in Automated Vehicles Using Multistage Attention-Based Convolutional Neural Network. IEEE Trans. Intell. Transp. Syst. 2020, 22, 4291–4300.
- 48. Zhang, P.; Xue, J.; Lan, C.; Zeng, W.; Gao, Z.; Zheng, N. EleAtt-RNN: Adding Attentiveness to Neurons in Recurrent Neural Networks. IEEE Trans. Image Process. 2019, 29, 1061–1073.
- 49. Krizhevsky, A.; Nair, V.; Hinton, G. CIFAR-10 and CIFAR-100 Datasets. 2009. Available online: https://www.cs.toronto.edu/~kriz/cifar.html (accessed on 8 August 2022).
- Jiang, H.; Tang, S.; Liu, W.; Zhang, Y. Deep learning for COVID-19 chest CT (computed tomography) image analysis: A lesson from lung cancer. Comput. Struct. Biotechnol. J. 2021, 19, 1391–1399.
- Lv, N.; Ma, H.; Chen, C.; Pei, Q.; Zhou, Y.; Xiao, F.; Li, J. Remote Sensing Data Augmentation through Adversarial Training. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2021, 14, 9318– 9333.
- 52. Ruhang, X. Efficient clustering for aggregate loads: An unsupervised pretraining based method. Energy 2020, 210, 118617.
- Riviere, M.; Joulin, A.; Mazare, P.-E.; Dupoux, E. Unsupervised Pretraining Transfers Well Across Languages. In Proceedings of the ICASSP 2020—2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Barcelona, Spain, 4–8 May 2020; pp. 7414– 7418.
- 54. Salur, M.U.; Aydin, I. A Novel Hybrid Deep Learning Model for Sentiment Classification. IEEE Access 2020, 8, 58080–58093.
- 55. Lu, T.; Du, Y.; Ouyang, L.; Chen, Q.; Wang, X. Android Malware Detection Based on a Hybrid Deep Learning Model. Secur. Commun. Netw. 2020, 2020, 8863617.
- 56. Basit, A.; Zafar, M.; Liu, X.; Javed, A.R.; Jalil, Z.; Kifayat, K. A comprehensive survey of Alenabled phishing attacks detection techniques. Telecommun. Syst. 2020, 76, 139–154.
- Fang, J.; Sun, Y.; Zhang, Q.; Peng, K.; Li, Y.; Liu, W.; Wang, X. FNA++: Fast Network Adaptation via Parameter Remapping and Architecture Search. IEEE Trans. Pattern Anal. Mach. Intell. 2020, 43, 2990–3004.
- 58. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Yu, P.S. A Comprehensive Survey on Graph Neural Networks. IEEE Trans. Neural Netw. Learn. Syst. 2021, 32, 4–24.
- 59. Huang, G.; Sun, Y.; Liu, Z.; Sedra, D.; Weinberger, K.Q. Deep networks with stochastic depth. In Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics); Springer: Cham, Switzerland, 2016; Volume 9908.

- 60. Chen, D.; Zhang, W.; Xu, X.; Xing, X. Deep networks with stochastic depth for acoustic modelling. In Proceedings of the 2016 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA), Jeju, Korea, 13–16 December 2016.
- 61. Koonce, B. SqueezeNet. In Convolutional Neural Networks with Swift for Tensorflow; Apress: Berkeley, CA, USA, 2021.
- 62. Bobenko, A.I.; Lutz, C.O.R.; Pottmann, H.; Techter, J. Checkerboard Incircular Nets. In SpringerBriefs in Mathematics; Springer: Cham, Switzerland, 2021.
- 63. Wang, S.; Zha, Y.; Li, W.; Wu, Q.; Li, X.; Niu, M.; Wang, M.; Qiu, X.; Li, H.; Yu, H.; et al. A fully automatic deep learning system for COVID-19 diagnostic and prognostic analysis. Eur. Respir. J. 2020, 56, 2000775.
- 64. Kumar, D.; Taylor, G.W.; Wong, A. Opening the Black Box of Financial AI with CLEAR-Trade: A CLass-Enhanced Attentive Response Approach for Explaining and Visualizing Deep Learning-Driven Stock Market Prediction. J. Comput. Vis. Imaging Syst. 2017, 3.
- 65. Cheng, X.; Zhang, Y.; Chen, Y.; Wu, Y.; Yue, Y. Pest identification via deep residual learning in complex background. Comput. Electron. Agric. 2017, 141, 351–356.
- He, S.; Jonsson, E.; Mader, C.A.; Martins, J.R.R.A. Aerodynamic Shape Optimization with Time Spectral Flutter Adjoint. In Proceedings of the AIAA Scitech 2019 Forum, San Diego, CA, USA, 7– 11 January 2019.

Retrieved from https://encyclopedia.pub/entry/history/show/66667