IMAGE FILTERING AND PROCESSING USING ADVERSARIAL NETWORKS

[1]MOHAMMED YASER ALI
[1]IV year, B.E. Department of CSE, BMS College of Engineering

[2]SIMRAN CHUGH
[2]IV year, B.E. Department of CSE. BMS College of Engineering

[3]Dr.G.R PRASAD
[3]Assistant Professor Department of CSE,BMS College of Engineering

[4]SANDEEP SURESH KUMAR YADAV
[4]IV year, B.E. Department of CSE, BMS College of Engineering

[5]KARTIK HAKIM
[5]IV year, B.E. Department of CSE, BMS College of Engineering

ABSTRACT: ***The promise of deep learning [5] is to discover rich, hierarchical models [1] that represent probability distributions over the kinds of data encountered in ARTIFICIAL INTELLIGENCE applications, such as natural images. Deep Generative models had been less impact due to difficulty of approximating many intractable probabilistic computations that arises in maximum likelihood estimation and related strategies and due to difficulties of leveraging the benefits of piecewise linear units in generative context. This Paper elaborates the uses of GANS to overcome these difficulties by filtering images of the real world faces under the Generative models.***

General Terms: Artificial Intelligence, Deep Learning, Neural Networks, LFW face image DATASET, conditions, performance Measures, Distribution probabilities

Keywords: GANS, Generator, Discriminator, Adversarial, Multilayer Perceptron's, Noise Images, Image processing [20].

1. INTRODUCTION

Learning of the Facial images has been potentially making life easier for all of us. Learning the facial image is changing the way we use and interact with the technologies. Some of them includes in Gaming, Tagging of persons, in the augmented Reality, Security and Many more.

Fig 1 - Facial image from LFW dataset

Many of the times the images are disturbed by the noise [3] . We may define the noise as the degradation in the image signal, caused by distortion. If an image has been sent electronically from one place to another we may expect the errors to occur in the image signal.

With advent of technology and success in artificial intelligence, there have been cases of image filtering where images are filtered to extent and improved clarity by automated programs, bots. Thus various forms of filtering have come into the picture. There are stochastic based, RBM's [4], back propagation based – all of which have their own technical drawbacks.

Deep Learning, with their remarkable ability to derive meaning from complicated or imprecise data, can be used to extract patterns that are too complex to be noticed by either humans or other computer techniques. A trained neural network [6] can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used for the noised images to answer to its best leaving the other technologies behind.

The Recent advancement in the neural networks is the Generative Adversarial networks (GAN'S) introduced by Ian GoodFellow in 2014[7]. This GAN's are compose of two distinct players A Generator [1] and the Discriminator [8]. These two are adversarial to each other i.e. the discriminator is tasked with distinguishing the samples from the model and samples from the training data; at the same time the generator is tasked with maximally confusing the discriminator. We can state the objective using the equation.

<u>The GAN's have to be trained for its performance. The training proposed in our GAN's is as follows.</u>

1. The generator outputs random RGB noise by default.
2. The discriminator learns basic filters in order to distinguish between face images and random noise.
3. The generator learns the skin tone and basic filters to confuse the discriminator.
4. The discriminator becomes more attuned to real facial features in order to distinguish between the tricky images from the generator and real face images. Furthermore, the discriminator learns to use signals in the conditional data to look for particular triggers in the image.

Advantage of Using GAN's over other types is that they can represent very sharp, even degenerate distributions [9]. GANs are a good method for training classifiers in a semi-supervised way. GAN's also gain some statistical advantage from the generator network not being updated directly with data examples, but only with parameters flowing through the discriminator. This means that components of the input are not copied directly into the generator's parameters.

This Literature Survey Paper intends to shed light on some of the Noise reduction techniques and their drawbacks. It also Discusses GAN's as an alternative and explores the existing mechanisms.

2. LITERATURE SURVEY

A survey of usability features for Noise reduction evaluates the currently available techniques of Noise reduction and tries to determine which of the existing mechanisms are reliable as well as usable. The authors state that with the increase in reliability, the speed of the mechanisms decreases. They also described the several types of existing techniques which exist in the world today.

Adaptive Wiener filters [10] are which is the optimal estimator (in the sense of mean squared error (MSE)) for stationary Gaussian process.. The Wiener filter doesn't tell you how to estimate statistics, it assumes you have the relations or spectrums. You can then search for different schemes to do the estimation. Wiener filter can be causal or non-causal, arbitrary ideal filters, or finite impulse response filters, discrete or continuous. It can be used for vector valued stochastic processes. A fixed filter will be less expensive computationally if there is not much change in signals. It will also be immune to estimation variance; whenever you estimate a stochastic parameter you will add noise to the system. Wiener also is not of much help when one have to deal with discrete state systems and with mixture distributions (such as Latent Dirichlet Allocation, K-means, and Nonnegative Matrix Factorization can attack).[11].

Median Filters are the technology where the Image if colored is converted to the gray scale image. The median of the pixels is found by sorting all the values in increasing order. Then the center pixel value is replaced with the median value. The impulsive noise is removed and then the result is passed to remove the blurredness and noise from the image. the median filter[6] considers each pixel in the image in turn and looks at its nearby neighbors to decide whether or not it is representative of its surroundings. Instead of simply replacing the pixel value with the mean of neighboring pixel values, it replaces it with the median of those values.

The median is calculated by first sorting all the pixel values from the surrounding neighborhood into numerical order and then replacing the pixel being considered with the middle pixel value. (If the neighborhood under consideration contains an even number of pixels, the average of the two middle pixel values is used.)

Anisotropic diffusion [13] Anisotropic diffusion: Perona and Malik anisotropic diffusion (PMAD) filtering, being the most common non-linear technique, was inspired from the heat diffusion equation by introducing a diffusion function that was dependent on the norm of the gradient of the image. The diffusion function, therefore, had the effect of reducing the diffusion for high gradients. It has been widely used for various applications such as satellite images, astronomical images, medical images, and forensic images. However, the PMAD filter has had two limitations up to now. First, it smoothes the information identically in all directions (isotropic). Second, the choice of the threshold on the norm of the gradient needed for the diffusion function is not evident, which makes the technique not truly automatic and cannot effectively remove noise [24]. Difficult to tune the decay of the distance function in anisotropic diffusion functional.

A **bilateral filter [15]** is a non-linear, edge-preserving and noise-reducing smoothing filter for images. The intensity value at each pixel in an image is replaced by a weighted average of intensity values from nearby pixels. This weight can be based on a Gaussian distribution. Crucially, the weights depend not only on Euclidean distance of pixels, but also on the radiometric differences (e.g. range differences, such as color intensity, depth distance, etc.). This preserves sharp edges by systematically looping through each pixel and adjusting weights to the adjacent pixels accordingly. Disadvantages include Staircase effect - intensity plateaus that lead to images appearing like cartoons. Gradient reversal-introduction of false edges in the image. When bilateral neighborhood size gets large then bilateral filtering is slow.

Restricted Boltzmann Machines [16] and neural networks in general, work by updating the states of some neurons given the states of others, so let's talk about how the states of individual units change. A restricted Boltzmann machine (RBM) a generative is stochastic artificial neural network[18] that can learn a probability distribution over its set of inputs. Each visible unit is connected to all the hidden units (this connection is undirected, so each hidden unit is also connected to all the visible units), and the bias unit is connected to all the visible units and all the hidden units. To make learning easier, we restrict the network so that no visible unit is connected to any other visible unit and no hidden unit is connected to any other hidden unit. The primary disadvantage is that RBMs are tricky to train well, since the common algorithm used, Contrastive Divergence, requires sampling from a Monte Carlo Markov Chain, and as such requires a bit of care to get things just right. Furthermore, that tricky partition constant in the model's energy makes computing the log likelihood under the model infeasible (or rather intractable), and so we cannot even track the loss we care about (let alone take derivatives with respect to).

Auto-encoders [19] learn an encoder function from input to representation and a decoder function back from representation to input space, such that the reconstruction (composition of encoder and decoder) is good for training examples. Regularized auto-encoders also involve some form of regularization that prevents the auto-encoder from simply learning the identity function, so that reconstruction error will be low at training examples (and hopefully at test examples) but high in general. Different variants of auto-encoders and sparse coding have been, along with RBMs, among the most successful building blocks in recent research in deep learning (Bengio et al., 2013b). Whereas the usefulness of auto-encoder variants as feature learners for supervised learning can directly be assessed by performing supervised learning experiments with unsupervised pre-training, what has remained until recently rather unclear is the interpretation of these algorithms in the context of pure unsupervised learning, as devices to capture the salient structure of the input data distribution. Auto encoders require huge pre-training techniques: it is greedy, i.e., it does not try to tune the lower layers in a way that will make the work of higher layers easier hence the cost of training is very high. The speed of the auto encoders is slow and they have the fuzzy design decisions. Auto encoders also lack in theoretical justification.

Ian Goodfellow Introduced Generative adversarial network which changed the revolution in the image processing. He trained adversarial nets on the Toronto Face Database (TFD) [23], He used noise as the input to only the bottommost layer of the generator network. Fig 2 shows Generative adversarial nets trained by him through simultaneously updating the discriminative distribution (blue, dashed line) so that it discriminates between samples from the data generating distribution (black, dotted line) from those of the generative distribution (green, solid line). The lower horizontal line is the domain from which noise is sampled, in this case uniformly. The horizontal line above is part of the domain of samples. The upward arrows show how the mapping imposes the non-uniform distribution on transformed samples. Generator contracts in regions of high density (peeks in the green) and expands in regions of low density.

Second-to-last column; this demonstrates that the generator has not learned to simply over fit the training data. Approach generates images using an iterative forward diffusion process (Sohl-Dickstein et al., 2015).

A cGAN could easily accept a multimodal embedding as conditional input .This input could be produced by a neural language model, allowing us to generate images from spoken or written descriptions of their content.

Soumith Chantala argued that parametric models for generating images have been explored extensively. However,

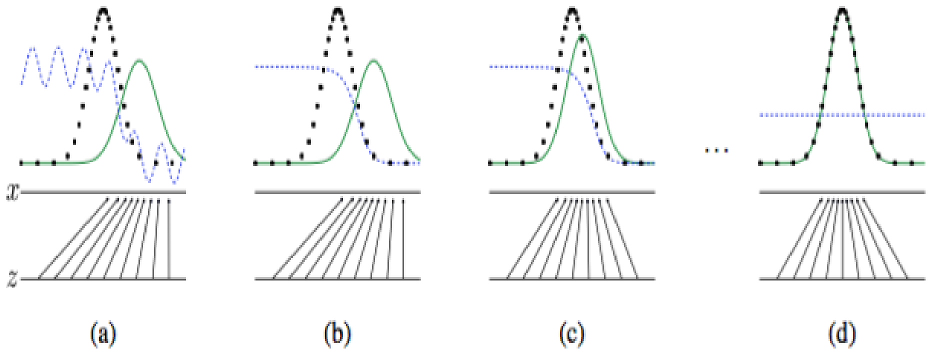

Fig 2– Training from GANS

Initially (in fig.a) the model is in the starting sate where the discriminator has deflection. But as the generation produces more and more samples ,The discriminator are prone to the noises and at the final steps of training if Generator and Discriminator have enough capacity, they will reach a point at which both cannot improve because The discriminator is unable to differentiate between the two distributions , i.e. D(x) =1/2

Jon Gauthier proposed a straightforward extension of the generative adversarial networks **called Conditional GAN's (cGAN's)[21]** adding a conditioning ability to the framework. We can establish some arbitrary condition for generation, which restricts both the generator in its output and the discriminator in its expected input. We might think of this condition as engaging both the generator and discriminator in a particular mode of generation or prediction. They applied this model to the facial dataset to deterministically control the attributes of the faces sampled from the model. **Figure 3** shows samples from a cGAN learned on this data. Each row begins with a conditional data value sampled from the empirical distribution. Each column is applied with the random shifts and samples an output image. The shifts cause noticeable differences in facial features, though it is evident that the slightly shifted values of y still lead to similar-looking faces.

The outlined face images at the far right show the nearest neighbor in the training data to the sampled image in the Generating natural images of the real world have had not much success until recently. A variational sampling approach to generating images has had some success, but the samples often suffer from being blurry.

In his proposed paper he declared the GAN's are unstable to train, often resulting the generators that produce nonsensical outputs. Hence they introduced the new architecture called **DCGAN [22]**. Then the architecture visualizes the filters learned by GANS and empirically show that specific filters have learned to draw specific objects He modified the model for the non-parameter learning using Convolution Net which is of parameter learning. He replaces the pooling layer of the convolutional networks with discriminators and Generators. They then applied the batch Normalization to both generator and discriminator to have the inputs with zero mean and unit variance. This helps them to deal with training problems that arise due to poor initialization and helps gradient flow in deeper models .Also the authors uses ReLU in generators for all layers except for the output . The author trained DCGANs on three datasets, Large-scale Scene Understanding (LSUN) [17] (Yu et al., 2015), Imagenet-1k and a newly assembled Faces dataset. No pre-processing was applied to training images besides scaling to the range of the tanh activation function [-1, 1]. Analysis has shown that there is a direct link between how fast models learn and their generalization performance (Hardt et al., 2015). The authors had the certain flaws and stated that further improvements could be made by fine tuning the discriminators representations.

Fig 3– face image filtering using cGAN

Stephen Koo proposed using adversarial framework by Ian Goodfellow the feed forward convolutional neural networks called DCGANS as discussed in the previous papers. He proposed to tackle the automatic colorization of black and white photos to combat the tendency of previous Coloring techniques.

He argued that the previous techniques use the Euclidean distance loss function which takes middle of the road solution not giving the appropriate outcome. He uses heuristics adapted from Larsen and Sønderby: for each iteration of training, if the discriminator's cross entropy loss is less than a certain margin, and then we skip the gradient update for Discriminator. This gives the generator a chance to catch up and harness the gradients from the discriminator before the discriminator starts to perform too well. We also transfer weights from the pre-trained baseline model to initialize the generator model: from this point, the generator need only learn how to employ the stochasticity of the noise input and continue optimize its outputs to fool the discriminator. This should give the generator an additional handicap to stay at pace with the discriminator.

3. CONCLUSION

In this paper, a detailed analysis of the image has been performed, highlighting the various social and technical shortcomings of the present methodology. It was seen how various techniques have been bypassed unfairly, leading to unfair outcomes. The requirement for newer technologies has led to thoughts about using adversarial networks rather than human working and older technologies capabilities, to process image processing.

Adversarial nets, the extended neural nets are thus can be used for a greater extent. The best training algorithms of adversarial nets are yet to be fully exploited yet it seems to provide a proper solution to the image filtering. The applications and shortcomings of various training techniques have also been discussed in the paper.

A Comparison of all the adversarial techniques mentioned in the paper revealed that the deriving the architecture from the convolutional nets and adding batch normalization seems to be highly effective and cost efficient. The only possible drawback seems to be the fact that fine tuning of the discriminator and training of the data, which can be overcome during implementation

4. ACKNOWLEDGMENTS

The work reported in this paper is supported by the college Through the TECHNICAL EDUCATION QUALITY IMPROVEMENT PROGRAMME [TEQIP-II] of the MHRD, Government of India.

5. REFERENCES

[1] C. M. Bishop and J. Lasserre, Generative or Discriminative? Getting the best of both worlds. In Bayesian Statistics 8, Bernardo, J. M. et al. (Eds), Oxford University Press. 3–23, 2007.

[2] Bengio, Y. (2009). Learning deep architectures for AI. Now Publishers.

[3] Philippe Cattin (2012-04-24). "Image Restoration:Introduction to Signal and Image Processing". MIAC, University of Basel. Retrieved 11 October 2013.

[4] Arochelle, H.; Bengio, Y. (2008). Classification using discriminative restricted Boltzmann machines (PDF). Proceedings of the 25th international conference on Machine learning - ICML '08. p. 536. doi:10.1145/1390156.1390224. ISBN 9781605582054.

[5] Schmidhuber., "Learning complex, extended sequences using the principle of history compression," Neural Computation, 4, pp. 234–242, 1992.

[6] D. C. Ciresan, U. Meier, J. Masci, J. Schmidhuber. Multi-Column Deep Neural Network for Traffic Sign Classification. Neural Networks, 2012.

[7] Goodfellow, Ian J.; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). "Generative Adversarial Networks". ArXiv:1406.2661

[8] McLachlan, G. J. (2004). Discriminant Analysis and Statistical Pattern Recognition. Wiley Interscience. ISBN 0-471-69115-1. MR 1190469.

[9] Bishop: Pattern Recognition and Machine Learning, Springer, ISBN 0-387-31073-8

[10] Haykin, Simon (2002). Adaptive Filter Theory. Prentice Hall. ISBN 0-13-048434-2.

[11] Florian Schmitt and Franz Rothlauf, "On the Mean of the Second Largest Eigenvalue on the Convergence Rate of Genetic Algorithms", Working Paper 1/2001, Working Papers in Information Systems, 2001.

[12] Huang, G. Yang, and G. Tang, "A fast two-dimensional median filtering algorithm", IEEE Trans. Acoust., Speech, Signal Processing, vol. 27, no. 1, pp. 13–18, 1979.

[13] Pietro Perona and Jitendra Malik (November 1987). "Scale-space and edge detection using anisotropic diffusion". Proceedings of IEEE Computer Society Workshop on Computer Vision,. pp. 16–22.

[14] R.A. Haddad and A.N. Akansu, "A Class of Fast Gaussian Binomial Filters for Speech and Image Processing," IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 39, pp 723-727, March 1991

[15] Carlo Tomasi and Roberto Manduchi, "Bilateral filtering for gray and color images," in Computer Vision, 1998. Sixth International Conference on. IEEE, 1998, pp. 839– 846.

[16] Larochelle, H.; Bengio, Y. (2008). Classification using discriminative restricted Boltzmann machines (PDF). Proceedings of the 25th international conference on Machine learning - ICML '08. p. 536. doi:10.1145/1390156.1390224.

[17] Yuan, Jiangye, Shaun S. Gleason, and Anil M. Cheriyadat. "Systematic benchmarking of aerial image segmentation." Geoscience and Remote Sensing Letters, IEEE 10.6 (2013): 1527–1531.

[18] Bottou, Léon (2004), "Stochastic Learning", Advanced Lectures on Machine Learning, LNAI, 3176, Springer, pp. 146–168, ISBN 978-3-540-23122-6.

[19] Modeling word perception using the Elman network, Liou, C.-Y., Huang, J.-C. and Yang, W.-C., Neurocomputing, Volume 71, 3150–3157 (2008), doi:10.1016/j.neucom.2008.04.030

[20] Rafael C. Gonzalez; Richard E. Woods (2008). Digital Image Processing. Prentice Hall. pp. 1–3. ISBN 978-0-13-168728-8.

[21] Learning (cs.LG); Artificial intelligence (sc.AI);computer Vision and Pattern Recognition (cs.CV); Machine Learning(stat.ML) Cite as: arXiv:1411.1784 [cs.LG]

[22] Learning (sc.LG); Computer Vision and Pattern Recognition (sc.CV) Cite as: arXiv:1511.06434 [cs.LG]

[23] Susskind, J., Anderson, A., and Hinton, G. E. (2010). The Toronto face dataset. Technical Report UTML TR 2010-001, U. Toronto.