COS 791 Assignment 1 Report

Joy Bengu

u25000307

***Abstract*—This attempt at assignment one plans to investigate how using meta-heuristic algorithms for the random creation and application of an image enhancement pipeline can improve the contrast and clarity of an image. Genetic Algorithm (GA) and Variable Neighbourhood Search(VNS) are the two algorithms utilised to evolve different image enhancement pipelines that consists of Gamma Correction, Unsharp Masking, Histogram Equalisation, Contrast Stretching and Gaussian Blur. The created pipelines were then optimised through the use of a combined fitness based on Mean Squared Error, Structural Similarity Index and Peak Signal-to-Noise Ratio and then compared the images that have been enhance with ground truth images. Results showed that GA converged faster and achieved higher fitness scores in training unlike VNS. The findings showed that meta-heuristics algorithms are effective when it comes to automatically enhancing an image thus not depending on the manual parameter tuning.**

***Index Terms*—image enhancement, contrast stretching, gamma correction, histogram equalization, unsharp masking, gaussian blur, variable neighbourhood search, genetic algorithm**

I. Introduction

Image enhancement is a crucial step in going towards the improvement of an image's visual quality for people who which to see the image and for automated analysis task (such as object detection and segmentation). The true tried-and-tested enhancement techniques such as gamma correction, unsharp masking and Gaussian blur depends manually choosing parameters which can be time-consuming and usually produce suboptimal results. This assignment looks into using meta-heuristic optimization method, more specifically Genetic Algorithms(GA) and Variable Neighbourhood Search (VNS), to automatically create and optimize image enhancement pipelines. The main aim is to find optimal sequences of contrast and brightness adjustments that gets as close to the ground truths as possible. The pipeline's performance is evaluated using Structural Similarity Index, Peak Signal-to-Noise Ration and Mean Square Error.
By following this approach, this assignment plans to demonstrate how meta-heuristic algorithms can automate the tuning and selection of parameters effectively thus improving the quality of the imaging without relying on traditional methods that having to be done manually

II. Background

A. *Image enhancement techniques*

Gamma Correction:

Gamma correction is an image enhancement technique used for the compensation of non-linear transfer characteristics of different input and output devices and to map out image intensities onto a unified intensity space. This allows for consistent and reproducible results, especially when the pixel values do not correspond with physical light quantities linearly.

Identify applicable funding agency here. If none, delete this.

The gamma value is a parameter that is usually between 1.8 and 2.5 and used in a power-law function to transform. Gamma values that below one will shift the image to be darker whereas gamma values above one will make the image brighter. Gamma values equal to one will have no effect on the image.

As [2] state, gamma correction is grounded on a power-law transformation. The formula is as follows:
$f_\gamma(a) = a^\gamma$
Where:

- a is the input intensity value
- f is the corrected intensity level
- γ is gamma which is the positive constant that will determine what shape will the transformation curve will take

The main properties of this function are:

- For values restricted to the interval of a [0,1], the value of f will stay in the range of [0,1] (essentially the function will always run through points[0,0] and [1,1])
- If γ=1, the identity function is f(a)
- If $\gamma < 1$, the curve will go above the diagonal, laying out a narrow range of dark input values going into a wide range of output values
- If $\gamma > 1$, the curve will go below the diagonal, laying out a wide range of dark input values going into a narrow range of output values

Gaussian Blur:

This is an image enhancement technique that is utilized in the smoothing or softening of an image by reducing noise and detail. It works as a low-pass spatial filter that applies normal distribution (a gaussian function) as a weighted average to the pixels[1][4] . Essentially every pixel will gain more weight according to the function and the weight decreases as the distance from the center pixel increases. It reduces high frequency components (sharp edges, noise,etc), thus making the image look soft and smooth in a way that resembles a translucent screen. Typical uses are enhancing background blur (to bring an object into focus), and obscuring details for privacy and security. It works in a following manner The gaussian function creates a bell-curve shape in which the pixel and weight are at their highest point and gradually lower further out The function is done in 2D to cover a kernel around the pixel The amount of blurring depends on the standard deviation of the gaussian distribution A higher sigma result will lead to more blurring spreading weight over more pixels Lower sigma results will results in a sharper image with less blur This is how the gaussian blur for 1D and 2D are defined: 1D Gaussian Function:

$$g_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}}\, e^{-\frac{x^2}{2\sigma^2}}$$

2D Gaussian Function

$$G(x, y; \sigma) = \frac{1}{2\pi\sigma^2}\, e^{-\frac{x^2+y^2}{2\sigma^2}}$$

x and y are the spatial coordinates relative to the kernel center (sigma) is the standard deviation which determines how much will the gaussian function will spread and in turn how much blurring is to be applied

Contrast Stretching:

As discussed by Gonzalez and Woods[5], contrast stretching allows for an image to cover the optimal full intensity range of a recording medium and /or display device by extending the range of intensity values. Low-contrast images, which are ideal for this technique, can happen due to serval reasons not limited to, insufficient lighting, limited dynamic range of the imaging sensor or improper lens aperture adjustment during the image capturing process.

It works with an intensity transformation function as show as:

$$s = T(r)$$

. where r is the input intensity and s is the output intensity at a given pixel. A piecewise linear transformation is a typical function used for this. It will be defined by two essential points (r_1, s_1) and (r_2, s_2)

The input image's intensity levels that lay within the range of $[r_1, r_2]$ are plotted onto a wider range of $[s_1, s_2]$ in the output image. Any r values that fall below a certain threshold of k are typically mapped to the values of s that are darker. Whereas the r values that are above k are mapped to s that are brighter. The function was made to have only one output for each input and to continually to rise ($[r_1 <= r_2$ and $s_1 <= s_2]$). This guarantees that the sequence of intensity levels screaming the same, bypassing any unwanted intensity distortions form emerging.

Unsharp Masking:

This is a technique that enhances the high-frequency details by subtracting a blurred version of an image. It is a way of increasing an image's contrast and and and thus making the image crisper. The process is as follows.

- The original image is blurred so that a smoothed version $f(x, y)$ is received
- A mask is made by subtracting the smoothed version from the original version $g_{mask}(x, y) = f_o(x, y) - f_b(x, y)$
- Add the weighted version of the mask to the original image to sharpen it $g(x, y) = f_o(x, y) + k * g_{mask}(x, y)$
  - $k >= 0$ controls the sharpening
  - $k = 1$ is standard unsharp masking
  - $k > 1$ is high-boost filtering which produces stronger sharpening
  - $k < 1$ the mask contribution is lowered for a gentler sharpening

Histogram Equalization:

Histogram equalization is an image enhancement technique that is used in the automatic adjustment of an image's intensity level so that the visual appearance and contrast is improved. The main goal of this technique is to locate an intensity mapping function that gives an output image with an approximately uniform histogram. Having the intensity values spread evenly allows for the image's contrast to be enhanced. Images that have histogram equalization applied visually become similar because their contrasts are vastly different. With that said, image may tend to look wither flat or muddy due to there not being string contrast differences that may be present in a non-uniform but visually attractive original

It is considered to be more of a fully automatic and hands-off approach, and, and it only relies on the information that was received from the image's histogram (which is cumulative or Cumulative Distribution Function [CDF]) without needing any parameters.

When dealing with digital images with discrete intensity levels, the transformation of $s_k = T(r_k) = (L - 1) * sum(p_r(r_j)) for j = 0 to k$. $p_r(r_j)$ is the intensity's normalized histogram component,r_j and L is the amount of possible intensity levels. The cumulative histogram $H(i)$ is calculated as the summation of all the histogram values $h(j)$ for $j <= i$

B. *Image enhancement and optimization*

1) *Genetic Algorithm for Image enhancement*

: As defined by Holland [3], genetic algorithms (GAs) are computer programs that develop and improve, similar to natural selection. These algorithms have the ability to address complex problems that may be difficult for their makers to fully comprehend. How these algorithms evolve fully depends on the two primary processes of natural selection and sexual production. Natural selection will choose which candidates within a population will be selected for reproduction. Sexual reproduction will mix and recombine genes among offspring. By doing this, evolution will be faster and there will be a wider area for exploration for potential solutions which may not be achieved by conventional programs.

Holland goes on to state that GAs can 'breed' programs that are able to solve problems even if their structure is not fully understood, thus showing their potential to be used in different complex systems. The key process of genetic algorithms are as follows:

1) Initialization: This is the creation of a population of probable solutions
2) Selection: Members within the population are evaluated based on their fitness. The fitness will determine which members are eligible for reproduction.

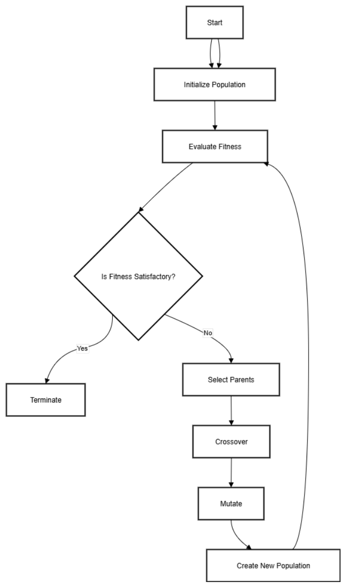

Fig. 1. Flowchart of a simple genetic algoirthm

3) Crossover: The genetic material from selected individuals are combined to create a new offspring (genes are allowed to be recombined)
4) Mutation: For the genetic diversity to be maintain and to explore potentially new solutions, at random, small changes are made to some of the offspring.
5) Replacement: Individuals with low fitness scores are replaced by the new offspring whilst keeping the initial size of the population.
6) Termination: The program either when a certain criteria is reached or when convergence is reached

2) Variable Neighbourhood Search for Image enhancement: Variable Neighbourhood Search (VNS) is a simple and effective meta-heuristic algorithm,developed by Mladenovic and Hansen.that was designed for the purpose for solving non-convex continuous and combinatorial optimization problems. VNS is different from other local search methods which use one neighbourhood structure and usually find themselves stuck in local optima with sub-par values. What VNS aims to do is to change a neighbourhood systematically within a local search algorithm.It will take take a look at growingly distant neighbourhoods of the initial solution and if it finds an an improvement it will cross ship to a new and better solution thus continuing its search from there. This allows for the favourable characteristics from the incumbent solution to be kept to create potentially promising solutions.

Mladenovic and Hansen went on to describe a simple VNS algorithm that works with a finite set of neighbourhood structures that were selected before, shown as N_k for $k = 1$, :k_{max}

1) Initialization: A set of neighbourhood structure that will be utilized in the search are chosen. A initial solution is found and denoted as x
2) Main Step:
   a) A neighbourhood counter is initialized as $k = 1$
   b) Until $k = k_{max}$ is reached the following is done:
      i) Shaking: A point x is generated at random the the k^{th} neighbourhood. This aids in avoid cycling.
      ii) Searching locally: A local search method is applied using x as the initial solution and x' is the local optimum that this search produced
      iii) Move or Increase
         A) if x' is better that the initial solution, x will now become the new solution and the search restarts to the first neighbourhood.
         B) Else if there was no improvement was found, the neighbourhood counter is increased by to explore a more distant neighbourhood in the following iteration
      iv) The main step will continue until a specific condition is reached.

III. Proposed Approach

A. *Genetic Algorithm for Image Enhancement Pipeline*

The GA was created for automatic optimization of the image enhancement pipeline by evolving the order of the enhancement techniques with randomly generated parameters. The approach started off with a population of candidate pipelines being initialization (varying in length and having a minimum length of 4 techniques) from these image enhancement techniques: gamma correction, gaussian blur, unsharp masking, histogram equalization and contrast stretching.
For every pipeline created, a fitness function that combined the three image quality metrics,comprising of Peak Signal-to-Noise Ratio, Structural Similarity Index and Mean Square Error, to show similarities between the enhanced image and the ground truth image. Parent selection was performed through the use of tournament selection, having a preference for pipelines with high fitness scores. New pipelines were created through crossover [which combined the genetic material of two parent pipeline] and mutation [parameters or techniques were at random altered for diversity]. The algorithm iterated until wither a certain amount of generations was reached or until convergence and the best pipeline was chosen based on fitness. The GA allow for an optimal sequence of enhancement techniques and parameter values to be discovered in tune to the training images.

B. *VNS for Image Enhancement Pipeline*

In an attempt to compliment the GA, VNS was implemented using systematic changes to the neighbourhood to avoid local optima when trying to optimize a pipeline. A pipeline was create at random and an iterative search was done by exploring that were growing distant to the current solution. Neighbourhoods were changed using the shaking methods, a technique was change within the pipeline, the existing

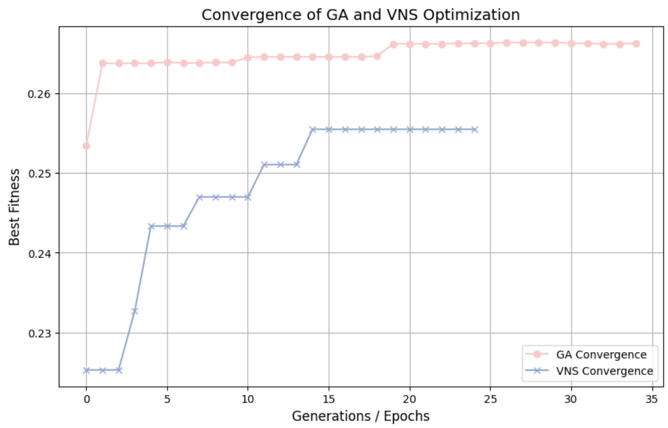

Fig. 2. Graph Showing Convergence

techniques had their parameters adjusted or a step was either removed or added in the pipeline. After this, a local search used mutation to further improve the changed pipeline over a fixed of times. If an improved pipeline was found in a neighbourhood, the search would start over from the first neighbourhood or else it would move to explore larger neighbourhoods. This process was done until either convergence was met or the maximum limit of epochs was reached. Evaluation of the fitness was the same as to the one used in the GA.

IV. Experimental Setup

A. Dataset Utilized

The assignment used a provided dataset of 10 training grayscale images and 5 test grayscale images, each image having a corresponding ground truth images.

B. Image Enhancement Pipeline

The pipeline consists of a sequence of a minimum of four techniques chosen from gamma correction, gaussian blur, histogram equalization, contrast stretching and unsharp masking . Parameter ranges for each technique are randomly initialized with set of predefined intervals [gamma in [0.5,2.0], amount in [0.5,2.0] and sigma in [0.5,5.0]].

C. Meta heuristic Algorithms

1) Genetic Algorithm:

- Population size: 30
- Generation Limit: 35
- Tournament Selection Size: 5
- Crossover Rate: 0.8
- Mutation Rate: 0.2
- Fitness Function: Weighted sum of PSNR, MSE ans SSIM between enhanced images and ground truth images

2) Variable Neighbourhood Search:

- Epoch Limit : 30
- Maximum neighbourhoods (k_{max}): 30
- Local search iteration per shake: 5
- Neighbourhood Modifications: Swapping of techniques, pipeline step addition or removal and parameter changes
- Fitness Function: Weighted sum of PSNR, MSE ans SSIM between enhanced images and ground truth

D. Fitness Evaluation

Fitness was calculated as a combination of Mean Squared Error (MSE) with a weight of 20 percent, Peak Signal -to-Noise Ratio (PSNR) and Structural Similarity (SSIM)weight a weigh tof 40 percent respectively. The value calculated was used to compare ground truth images to the images that were enhanced by the pipeline.

E. Implementation

For the implementation, Python 3 was used with the following libraries: scikit-image,OpenCV, NumPy and joblib (joblib was used for parallel fitness computation). Training pipelines were optimized on the images from the training and with the best results,the test images were tested.

F. Computational Resource

Experiments were conducted on Google Colab with with a T4 GPU with the latest version of runtime

V. Results

As the GA and VNS generate the pipelines at random, below are the results from a single run,

1) Training Data: : Below are the details of the best pipelines found during the run

Method	Step	Technique	Parameters
GA	1	Gamma Correction	$\gamma = 1.6128$
GA	2	Gaussian Blur	$\sigma = 2.9531$
GA	3	Gaussian Blur	$\sigma = 2.9531$
GA	4	Gaussian Blur	$\sigma = 2.4350$
GA	5	Gamma Correction	$\gamma = 0.5829$
Best GA Fitness Score			0.2664
VNS	1	Gamma Correction	$\gamma = 0.5061$
VNS	2	Gaussian Blur	$\sigma = 3.5585$
VNS	3	Gaussian Blur	$\sigma = 3.8284$
VNS	4	Gaussian Blur	$\sigma = 1.6501$
Best VNS Fitness Score			0.2555

TABLE I
Best GA and VNS Pipelines Found During Training with Fitness Scores

2) Testing Data: : Below are the details of the best pipelines being applied to the test images:

Image	GA			VNS		
	MSE	**PSNR**	**SSIM**	**MSE**	**PSNR**	**SSIM**
Image 1	1803.20	15.57	0.4564	9164.20	8.51	0.3481
Image 2	9637.38	8.29	0.1803	7113.84	9.61	0.2149
Image 3	8349.72	8.91	0.3746	4774.89	11.34	0.1178
Image 4	2733.77	13.76	0.2518	4134.14	11.97	0.2394
Image 5	4823.53	11.30	0.2474	8502.08	8.84	0.1385

TABLE II
Comparison of GA and VNS Test Results on Images (MSE, PSNR, SSIM)

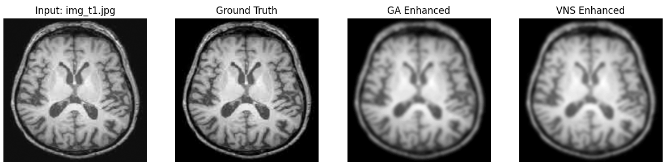

Fig. 3. Test Image: $img_t1.jpg$

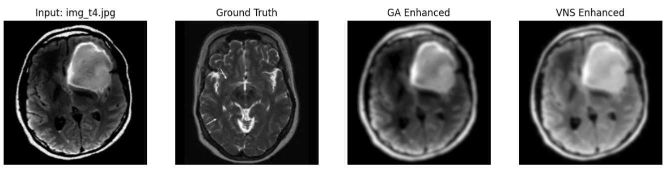

Fig. 4. Test Image$img_t4.jpg$

A. Discussion

1) Training results: GA and VNS were both employed in the training phase to optimize the created pipelines by adjusting the techniques' parameters (the ones that have parameters). The convergence braphs shows the progression of fitness over generations in GA and epochs for VNS. It displays how fitness steadily improved and then plateaued as algorithms reached convergence. The best pipeline found the GA was a combination of gamma correction and gaussian blur which achieved a fitness score of 0.2664 as compared to the best pipeline found by VNS, which was coincidentally, also a combination of gamma correction and gaussian blur with a fitness score of 0.2555. This shows that the GA was able to cover more of the search space effectively compared to VNS during training.

2) Test results: When looking at the the test set, both GA and VNS pipelines able to enhance the image quality when comparing to the ground truth images(as shown using the metrics of MSE, SSIM and PSNR). Across the five test images, GA achieved a much lower MSE and both higher PSNR and SSIM values when being compared to VNS, solidifying the performance achieved during training. Visually comparing the output images produced by GA (the enhancement done in GA are more balanced) and VNS shows that clarity and contrast relative to the ground truth images.

VI. Conclusion

This assignment was able to apply meta-heuristic algorithms, namely Genetic Algorithm and Variable Neighbourhood Search, to create pipelines meant for image enhancement. The GA was able to perform better than the VNS, as GA delivered higher fitness scores and better test image quality metrics. This shows that the image enhancement was able to get close to the ground truth. The results showed the potential of evolutionary algorithms and search methods to process images when traditional manual tuning may waste time.

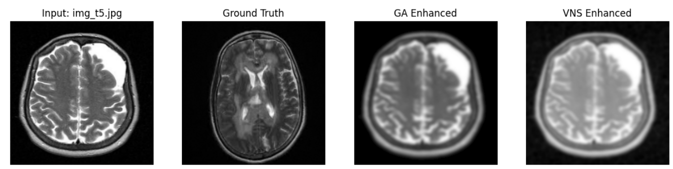

Fig. 5. Test Image $img_t5.jpg$

References

[1] Burger, W. & Burge, M. Digital Image Processing: An Algorithmic Introduction. (Springer Cham,2022), https://link.springer.com/book/10.1007/978-3-031-05744-1bibliographic-information

[2] Holland, J. Genetic Algorithms. , http://www.jstor.org/stable/24939139, Scientific American, a division of Nature America, Inc.

[3] Mladenović, N. & Hansen, P. Variable neighborhood search. *Computers Operations Research*. **24**, 1097-1100 (1997), https://www.sciencedirect.com/science/article/pii/S0305054897000312

[4] Szeliski, R. Computer Vision: Algorithms and Applications. (Springer International Publishing,2022), https://books.google.co.za/books?id=A34ZygEACAAJ

[5] Gonzalez, R. Digital Image Processing. (Pearson Education,2009), https://books.google.co.za/books?id=a62xQ2r$_f8wC$