Trinity College Dublin
Coláiste na Tríonóide, Baile Átha Cliath
The University of Dublin

School of Computer Science and Statistics
School of Engineering

Photograph Palette Alteration Utilising Image Segmentation

Daniel McCormack
12305446
Supervisor: Prof. Aljosa Smolic

A dissertation submitted in partial fulfilment of the degree of
M.A.I. (Computer Engineering)

Submitted to the University of Dublin, Trinity College, May 2018
Trinity College Dublin
University of Dublin

Declaration

I, Daniel McCormack, declare that the following dissertation, except where otherwise stated, is entirely my own work; that it has not previously been submitted as an exercise for a degree, either in Trinity College Dublin, or in any other University; and that the library may lend or copy it or any part thereof on request.

Signature: ______________________________

Date: ______________________________

Summary

This project will examine in detail the use of semantic segmentation for the purpose of colour manipulation in images. Previous work with palette alteration for image manipulation were utilised to this end. The aim of this combination of techniques was to provide a means for image manipulation for novices comparable to the standard achievable by experts. Examinations of the cutting edge are made, with the approaches that are at the current forefront of study in the area, as well as a novel approach to palette estimation to further simplify the proposed approach. Results are examined for each stage of the process, and final results are compared to similar methods of palette based image manipulation without the use of semantic segmentation.

Acknowledgements

I would like to thank Aljosa Smolic for providing me with the opportunity to work on this project, and meeting with me throughout. I also have to thank my family for their encouragement throughout the project. I would also be remiss not to thank my friends who have kept me company and helped me throughout the year, most notably among them Simon, for helping proof read segments of the report, and offering formatting advice.

I also have to express to greatest of thanks to Mairéad Grogan for her advice over the duration of the project. Without her and Matis Hudon advising me on how to best proceed with the research this project would not have been possible. Taking the time to meet with me regularly throughout the latter half of the project took time out of both of their days and Mairéad's previous work and familiarity with the field of palette manipulation helped see the project to its finality.

Photograph Palette Alteration Utilising Image Segmentation

by

Daniel McCormack

12305446

Submitted to the School of Computer Science and Statistics

Trinity College Dublin on 17th of May 2018

as part of the degree of M.A.I. Engineering

Abstract

This project will examine in detail the use of semantic segmentation for the purpose of colour manipulation in images, using previous work with palette alteration to further simplify the approach for those with little experience with image manipulation software. The aim of this combination of techniques will be to provide a means for image manipulation for novices that is near the standard achievable by experts. Examinations of the cutting edge will be made, with the approaches that are at the current forefront of study in the area, as well as a novel approach to palette estimation to further simplify the proposed approach. This approach was found to be good at what it aimed to do, though there is still work that could done with regards the novel palette estimation.

Table of Contents

Chapter 1:

Introduction

In this chapter I will give an account of the aims of the project and the motivations for these aims, as well as a breakdown of the structure of the report.

1.1 Motivations

Many attempts have been made to close the gap between novice and expert uses of photo manipulation software, particularly in relation to colour manipulation. Approaches have consisted of defining palettes for either the target photo, or for both the initial photo and the target photo, without consideration for wanting to alter an area of an image without altering another identically coloured area. Segmentation has been considered in the past, primarily with approaches relating to colour, or contiguous coloured regions, without consideration for the fact that borders are subtle and not always defined by hard edges. In some cases, hard edges are the opposite of what might be expected, due to gradients between regions being common, particularly in medical applications of computer vision technology.

Artificial Neural Networks are on the forefront of discussion in the field of computing in recent years, with the various Artificial Neural Network architectures finding use in stock predictions, to recommender algorithms on popular sites, to future autonomous vehicles. Organisations are making interesting developments in the field of computer vision and image manipulation through the use of these Artificial Neural Networks, particular with Deep Neural Networks and Convolutional Neural Networks that have been shown to particularly excel in the field of computer vision.

Taking inspiration from these two areas, I sought to explore a method which could create a fast and accessible method for dynamic and selective feature recolouring.

1.2 Aims

This project aims to investigate and analyse the use of feature based semantic segmentation for the editing of images, particularly for extracting and editing the palettes of images. With this, the goal is to allow for a more accessible method of recolouring regions of a photo for novices. The use of recent advances in semantic segmentation provide a new and interesting approach to the segmentation of these images, so use and current validity of one of the cutting edge of these approaches will be examined in detail.

For this purpose it was decided to examine DeepLab semantic image segmentation convolutional neural network to extract specific objects from an image. The extracted images with pixels assigned object values are then saved as new image and a novel palette extraction algorithm is used to identify the colour clusters present. Using this palette information and the extracted object image, the program created by Grogan 2017[1] was used to recolour the image, though any method with similar parameters could be applied. Finally, this image is recombined with the other extracted image segments, to create a final completed image with just the colour of one segment augmented. Alternatively, the multiple segments could be recoloured individually, to allow for even greater control and customisation.

1.3 Outline

Chapter 2 of this paper will detail the state of the art in relevant areas of study. Among these are semantic segmentation and palette based image manipulation approaches. I will also examine colour quantization, and by extension clustering of multidimensional data with the aim of examining the usefulness of these approaches for palette estimation.

Then with that knowledge Chapter 3 will go into detail the method being proposed step by step. Some samples of each step in operation will be shown, particularly with a 1-dimensional version of the 3-dimensional algorithm for palette estimation being shown for clarity on the approach.

Chapter 4 will examine the results from the proposed system at each step of the process. First segmentation, then from the segmentation, to the palette estimation, to the final results. These final results will then be compared to other methods for altering images by changing palettes.

Finally Chapter 5 will detail the pros and cons of this approach. Areas in which more study is needed will also be looked at, particularly focusing in perceived shortcomings of the solution and ways in which these might be fixed with further study.

Chapter 2:

State of the Art

2.1 Semantic Segmentation

The first step in performing this analysis was the segmentation itself. Semantic segmentation involves assigning each pixel in the image an 'object' value. This differs from typical image segmentation in that it attempts to assign meaning to each segment, rather than just attempting to connect contiguous regions of an image. Several approaches have been tried to achieve semantic segmentation, but recently the most promising results, with regards to both accuracy and versatility have been from fully convolutional networks (FCNs)[2]

These FCNs are trained on datasets of images such as the PASCAL datasets[3], these datasets have images which are segmented beforehand, and assigned labels. In the case of the 2011 Pascal dataset the labels are as follows:

Person: person

Animal: bird, cat, cow, dog, horse, sheep

Vehicle: aeroplane, bicycle, boat, bus, car, motorbike, train

Indoor: bottle, chair, dining table, potted plant, sofa, tv/monitor

The training images used include some combination of these objects, and examples are shown in figure 2.1

Fig 2.1: Segmentation map

The PASCAL VOC, the de facto standard for training and evaluation of new computer vision algorithms, performed challenges for the years of 2005 to 2012[4], but still keeps its' datasets and evaluation servers active. To accompany this, there is a leader board detailing the best evaluated methods of image classification and segmentation. Since the 9th of February 2018, and up to the date of writing, the current best evaluated method for image segmentation is Google Inc's Deeplabv3[5],with an accuracy of 89%[6].

DeepLab in particular uses a method of convolution and deconvolution to achieve more accurate semantic segmentation, as shown in the diagram below[7]:

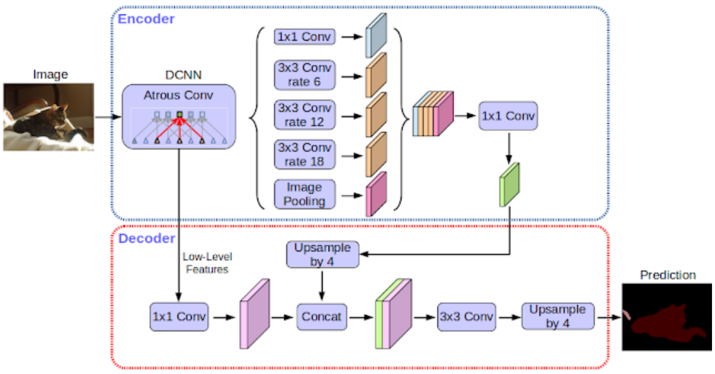


Fig 2.2: Flow of operation in DeepLab CNN

The particular contribution which DeepLab provides to the area of semantic segmentation is the use of atrous convolution and atrous spatial pyramid pooling to achieve dense object classification using Conditional Random Fields to combine the outputs of the network, as detailed in figure 2.3:

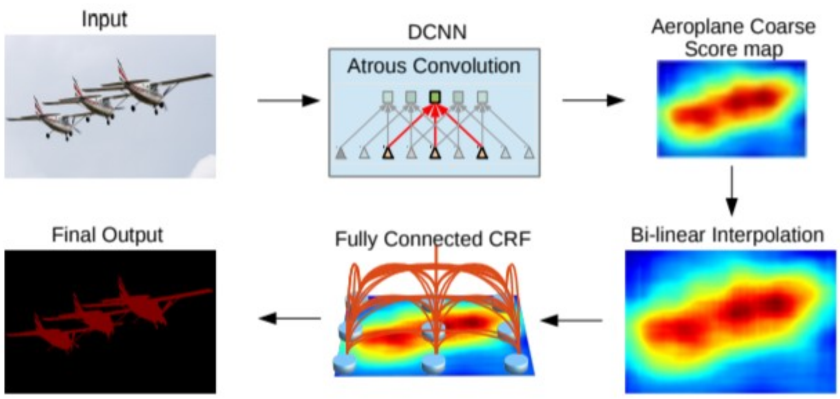


Fig 2.3: Impact of Conditional Random Fields for semantic segmentation

2.2 Colour Quantization

Fig 2.4: Image going through colour quantization (256, 32,16 then 8 colours)

The next step necessary for performing palette alterations was estimation of the palette itself. Colour quantization is an area of study typically applied to finding colours to represent an image using fewer colour entries (often 256, 64 or 16). Extracting a palette for an image is similar to this as a number of colour entries need to be selected which will be indicative of the whole image. As this paper will discuss a method not related to image formats or data limitations, the number of palette entries is not known beforehand, and the primary focus is on the appropriate number of colour entries to allow for it to closely match human perception.

Creating a palette for an image of arbitrary size involves clustering the pixels of the image into an appropriate number of clusters, and then using a value taken from these clusters to indicate the value of the palette entry. As most colour spaces are represented using a 3-dimensional coordinate space, any method used for clustering 3-dimensional data can be applied for colour quantization and vice versa. As these are discussed in this paper, it should be noted for clarity that the data points discussed refer to the pixel values of the image, and the clusters to the individual palette entries.

2.2.1 Methods for Clustering of 3-Dimensional Data

2.2.1.1 Median Cut

Median cut is a method for the clustering of data across an arbitrary number of dimensions by dividing the entirety of the data through the median along the longest of its dimensions[8][9]. This is then repeated across all of the separated sections until the desired number of colours is reached. This has been used for reducing the number of colours in images for displays and image formats with limitations in this regard. The main limitation of this approach is that by dividing each subregion at each step, this limits the number of palette entries to powers of two. For any other values some collation of entries need to be performed.

Fig 2.5: Palette estimation using median cut[10]

2.2.1.2 K-means clustering

K-means clustering[11] is one of the most well known methods for clustering of n-dimensional data. K-means clustering uses 'k' centroids which move with each iteration to the mean of the points which relate to those centroids. The centroid which relates to a data point is defined by proximity, and this causes the centroids to move to local minima as iteration occurs.

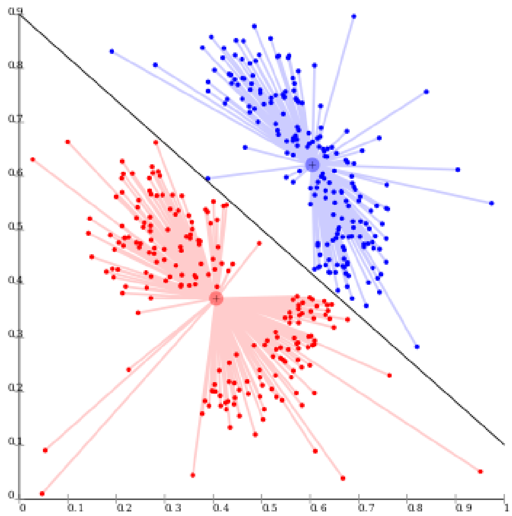


Fig 2.6: K-means clustering

K-means is frequently used for colour quantization, but it tends to find the mean value of the colours of the image rather than the most significant, and can fail if the data is too uniform, or if the incorrect value for k is selected. For example, if there are three clearly defined clusters, but the value for k is set to 4, one or more of the clusters will not be correctly represented.

2.2.1.3 DBSCAN

Density-based spatial clustering of applications with noise is a clustering approach by Ester et al.[12] that creates clusters based on how many nearby neighbours a point has, marking any points without nearby neighbours as outliers. In order to do this it requires two values, ε and the minimum number of points to qualify as a dense region. A point found is used as a core and any points within the region of ε are added and the process repeats until the entirety of a cluster is found. This process then repeats to find any remaining clusters. This provides the advantage of not needing a pre-defined number of clusters, like k-means or median cut. That said, while applicable to colour quantization, it may not be ideal for these applications, as the clusters found can be abnormally shaped, and finding a single value indicative of each cluster is not always possible. This is obvious from clustering shown in figure 2.6 below

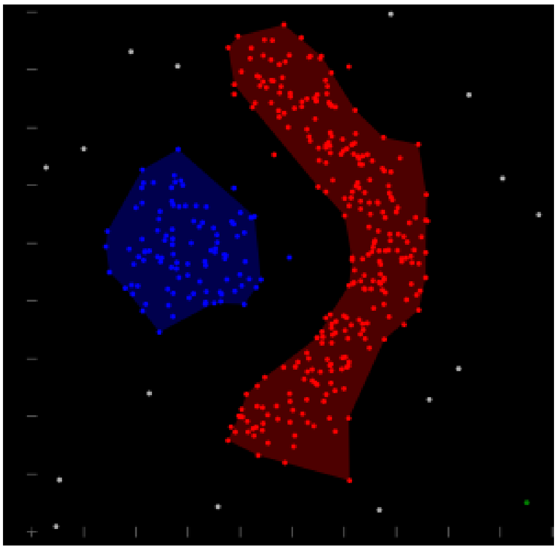


Fig 2.7: DBSCAN Clustering

2.2.1.4 OPTICS

Ordering points to identify the clustering structure is an algorithm similar to DBSCAN devised by Ankerst et al[13], which also takes into account the density of clusters. It also uses ε and a minimum points threshold to find clusters, but it stores each point linearly so that the closest points are neighbours in ordering. It also stores a distance for each point that represents density that needs to be accepted for a cluster to contain two distinct points. This is shown in the figure below, there the blue region is more densely packed than the red region, allowing the red region to capture some less dense points round the blue region.

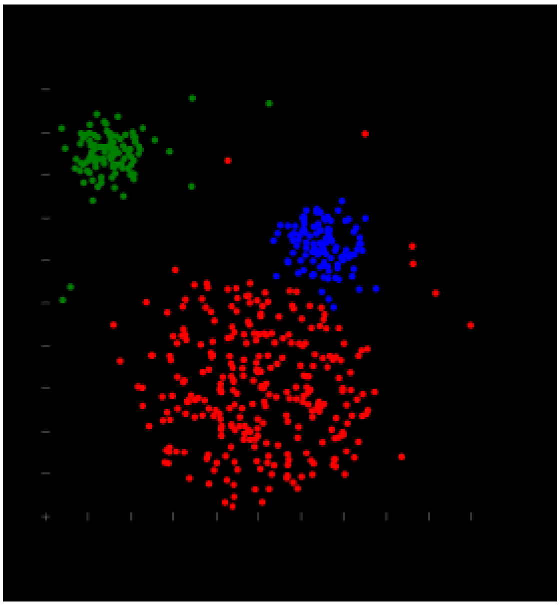


Fig 2.8: OPTICS Clustering

While this does help account for some of the limitations of the DBSCAN algorithm, it does not improve in any significant way to make it more suited to colour quantization.

2.2.2 Clustering methods applied to colour quantization

The methods discussed have positives and negatives for the application of colour quantization. Median cut and k-means address the problem more directly, as the key feature of colour quantization is not the continuity of the clusters found, but the centres of the clusters and their relation to their components. They are not ideal for the application of finding the palette of an image, as they require a pre-defined ‘k’ to define the number of clusters to be found. DBSCAN and OPTICS provide an interesting approach to dynamic identification of the number of clusters present in a data set, but the remainder of the clustering performed with these methods is not applicable to colour quantization, as these methods allow for abnormally shaped clusters to form, where the “centres” of these clusters may have little significance to the clusters themselves. For example, a ring could form around a second internal cluster, resulting in two clusters with identical “centres” making the density based clustering inappropriate when it allows for clusters of irregular shapes.

2.3 Palette Based Image Recolouring

A palette based approach to photo recolouring is often taken as a means to make editing software more accessible for novices while still allowing for creative applications of recolouring[14]. The primary appeal of these methods is the accessibility it provides while still allowing for drastic changes to the output image. These methods apply theory from colour transfer to allow for finer control of end results.

These methods have been shown to be more accessible for novices when compared to methods preferred by experts. That said, use of these methods often results in artefacts caused by regions of the image not intended to be altered to be manipulated along with the region for which the change was intended. i.e. Make a green t-shirt blue, but in doing so also make the bush behind the t-shirt blue.

This is an area of study that has seen some work, more recently using Colour Decomposition Optimization[15] and User specified L_2 values[1] to attempt to account for limitations in the approach, particularly in the artefacts it can create.

While the method proposed by Grogan et al.[1] in particular has been promising in reducing these issues, it is still not perfect, as adding too many user defined constraints to a palette alteration can result in the desired changes not taking place at all, or can create other, unexpected artefacts.

Chapter 3:

Proposed Method

The method proposed by this paper utilise the current state of the art for each step of processing the image. The image is first segmented before any alteration or palette estimation takes place. This is to more accurately allow for editing of specific objects in the scene, rather than global alterations. The segmented images then have their palette extracted using a novel method, designed to extract a suitable number of unique colours from the image, avoiding colours too similar to one another, while making sure to get a sufficient number of colours such that the palette of the images are representative of their contents. These segments are then altered as desired, by taking the input palette and replacing it with a new palette, using methods which have been show to be practical for these applications. These edited segments are then recombined into a new image representing the changes applied to the original image.

3.1 Semantic Segmentation

For the purpose of segmentation DeepLab will be used. This is because of the tested accuracy of the network for the purposes of semantic segmentation, as well as limitations in personal hardware for training of a segmentation network from scratch. The network will take an input of an image and assign an object value to each pixel in the image. The output of this step will be a mask for each type of object in the scene. These masks are used to extract regions of the image for each object type, as shown below:

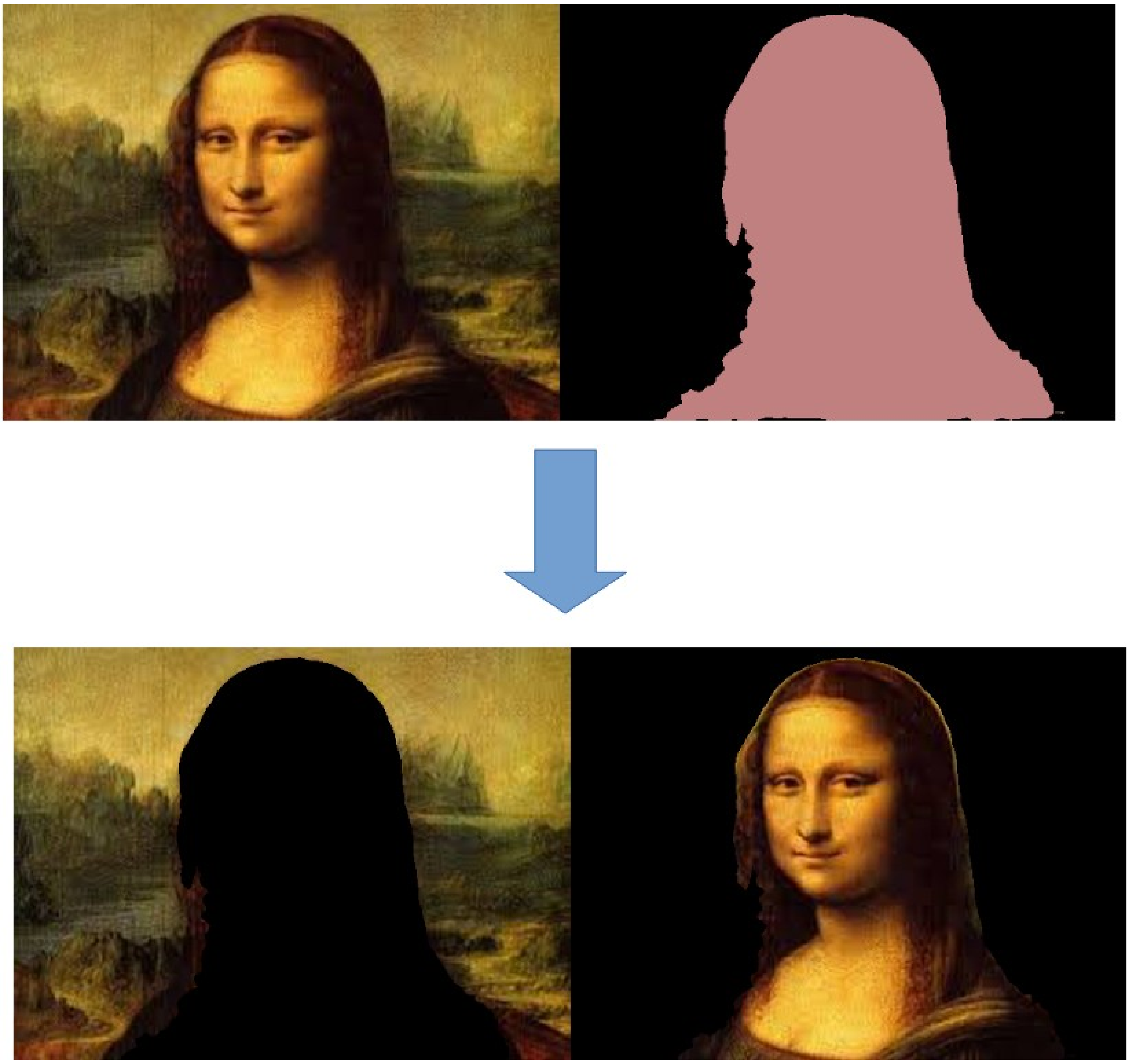


Fig 3.1: Segmentation mask applied to separate image by object value

3.2 Palette Estimation

The next step involves estimation of a palette for each of these regions. The method used is a novel greedy clustering algorithm. The primary goal of this new method is to present a palette with significantly varied entries to the user, such that there is little confusing over which palette entries relate to which parts of the image. As such a number of palette entries that is too high or too low will result in either being overly vague or overly specific. The ideal number of palette entries varies by image, and as such the method will need to vary to match. Similar to DBSCAN and OPTICS, rather than use a pre-determined k value, a radius(ε) and % threshold(θ) are defined instead.

The method devised for the estimation of palette was based on the approach discussed by Aksoy et al.[16] for palette estimation, though that approach was limited in that it assigned a voting mechanism to account for removing colours that were parts of gradients from the end, as the aim of that method was to find seed pixels for colour based segmentation rather than a palette to present to an end user.

The method does no such voting or discounting, but instead considers each pixel irrespective of it's position, and accounts for the number of non-zero pixels in the image. This allows for a dynamic number of palette entries, more suitable for human input, while allowing for smaller segments extracted by semantic segmentation to still have a suitable palette based on internal pixels, rather than accounting for the entire image. To begin with the image segment is broken into a 10 x 10 x 10 histogram, and the largest value is selected. If the sum of values within ε of the selected point are greater or equal to θ% of the image, then the selected region is broken into a new smaller 10 x 10 x 10 histogram, and the local maxima is selected as the palette entry. Then every entry on the original histogram within ε of the original point are set to zero, and the process is repeated until no further entries can be found.

An example of this in a 1-dimensional case is shown below, with the following values:

θ=25% (0.25)

ε=1

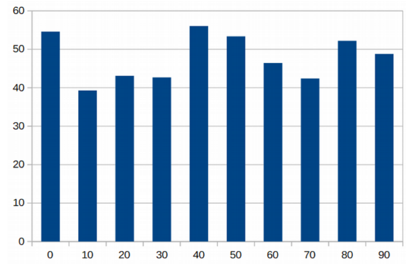


Fig 3.2: Randomly selected values, in a 10 bin histogram

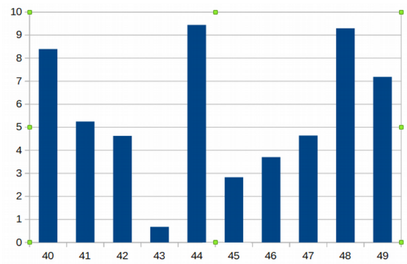


Fig 3.3: Local histogram to find optimal point value

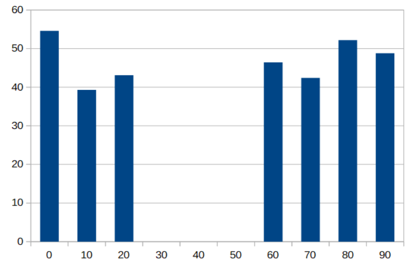


Fig 3.4: Histogram after exclusion of the region defined by ε is performed

This example results in the entry at 44 being used as a palette entry, before the process is repeated. Once a complete palette is reached it is returned to the user of to the next stage of the program. The limits of this approach are that there can never be more than θ^{-1} palette entries. The number of spherical regions which can be made in the space is also a limiting factor, as there cannot be more palette entries than spheres fit in the region. The mathematics of this sphere packing is outside the scope of this project, but some data points can be made. When the value of ε is 1, in 10 x 10 x 10 space, a full 1000 palette entries can be found. If ε is greater than 17.32 in the same space, only one can be found, as the spherical regions may overlap, but the centre of one sphere cannot exist within another. From this, any value less than or equal to 17.32 must allow for at least 2 entries. Also it should be noted that if a region of 0 was defined, hypothetically infinite palette entries might be found, so the formula representing the maximum number of entries most likely involves division by ε. With that in mind these values do result in interesting interactions as they are altered, making them much more significant than the sum of their parts.

Running this algorithm on the same image with different parameters the following table was generated of the number of palette entries. The image used was the background from Leonardo Da Vinci's Mona Lisa, shown below:

Fig 3.5: Mona Lisa (background)

Table 3.1: Area vs Threshold for palette estimation

	θ=0	θ=0.01	θ=0.02	θ=0.04	θ=0.08	θ=0.18	θ=0.32
ε=1	34	7	4	4	1	1	1
ε=2	9	4	3	3	2	1	1
ε=3	5	3	2	2	2	1	1
ε=4	4	2	2	2	2	1	1
ε=5	2	2	2	2	2	1	1
ε=6	2	2	2	2	1	1	1

As can be seen from the table above, generally as θ and ε increase the number of entries found decreases. It was from testing like this that ideal values were found for these input variables, of $\varepsilon=3$ and $\theta=0.01$, as this generally seems to avoid having too many overly similar entries, while still having a not insignificant density threshold for regions of the histogram which are to be used for palette entries. While this provides an accurate method of gathering a palette from an image when compared to methods in more common use like median cut or k-means, it is slightly more computationally expensive, taking longer to extract a full palette. When comparing these to other methods it must be considered that 1 to 1 comparison is not entirely accurate as when using median cut, the number of palette entries desired needs to be defined beforehand, so median cut does not entirely fill the needs of this approach.

3.3 Palette Alteration

At this step in the procedure user input is needed to specify which segment's palette needs to be altered, as well as which alterations need to take place. These alterations are performed on the relevant segments, and the results are then recombined into a singular image. This image is then returned as the output of the entire process. The primary method employed for this stage of the process is the use of the a method proposed by Grogan et al[1], with no user defined constraints. The results from this step will be shown in detail in the next chapter of this report.

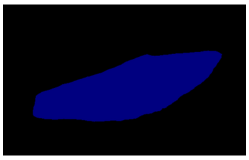
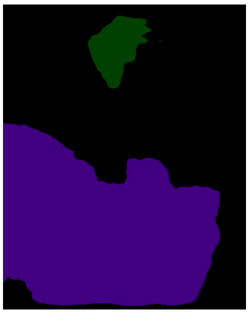
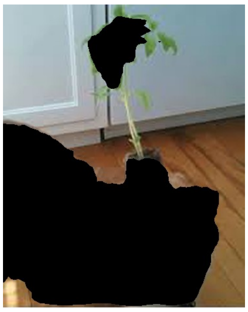
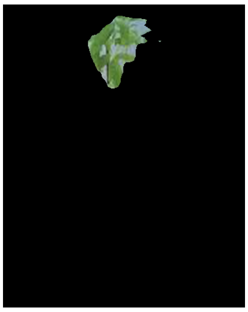
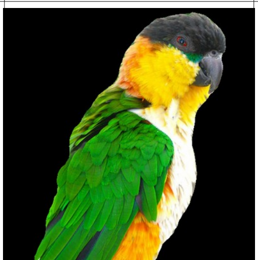
Chapter 4:

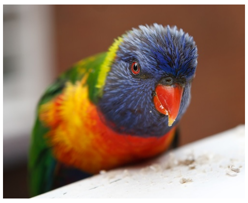
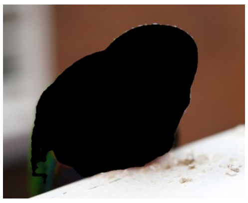
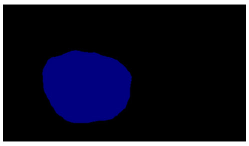
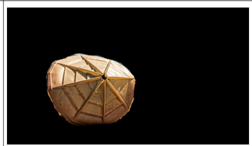
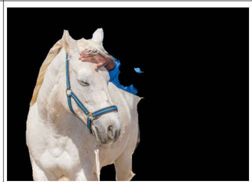
Results

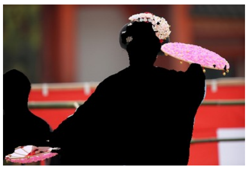
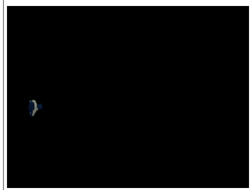
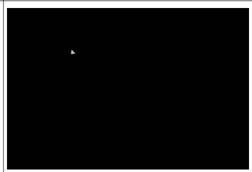
The results for this application are relatively subjective, and as a result it is difficult to quantify the results in any accurate manner. The results will be shown for sample images for both the segmentation and the palette extraction stages, as well as final images following alterations taking place. Over the next few pages these will be shown step by step. The images and components thereof will be referred to afterwards by their ID numbers, and segment numbers, with segment 0 being the entire image.

4.1 Segmentation

Table 4.1: Table of Semantic Image Segmentation with Image IDs

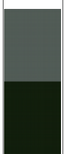
ID	Original Image	Segmentation mask	Segment 1	Segment 2	Segment 3
1					**N/A**
2					**N/A**
3					
4					**N/A**

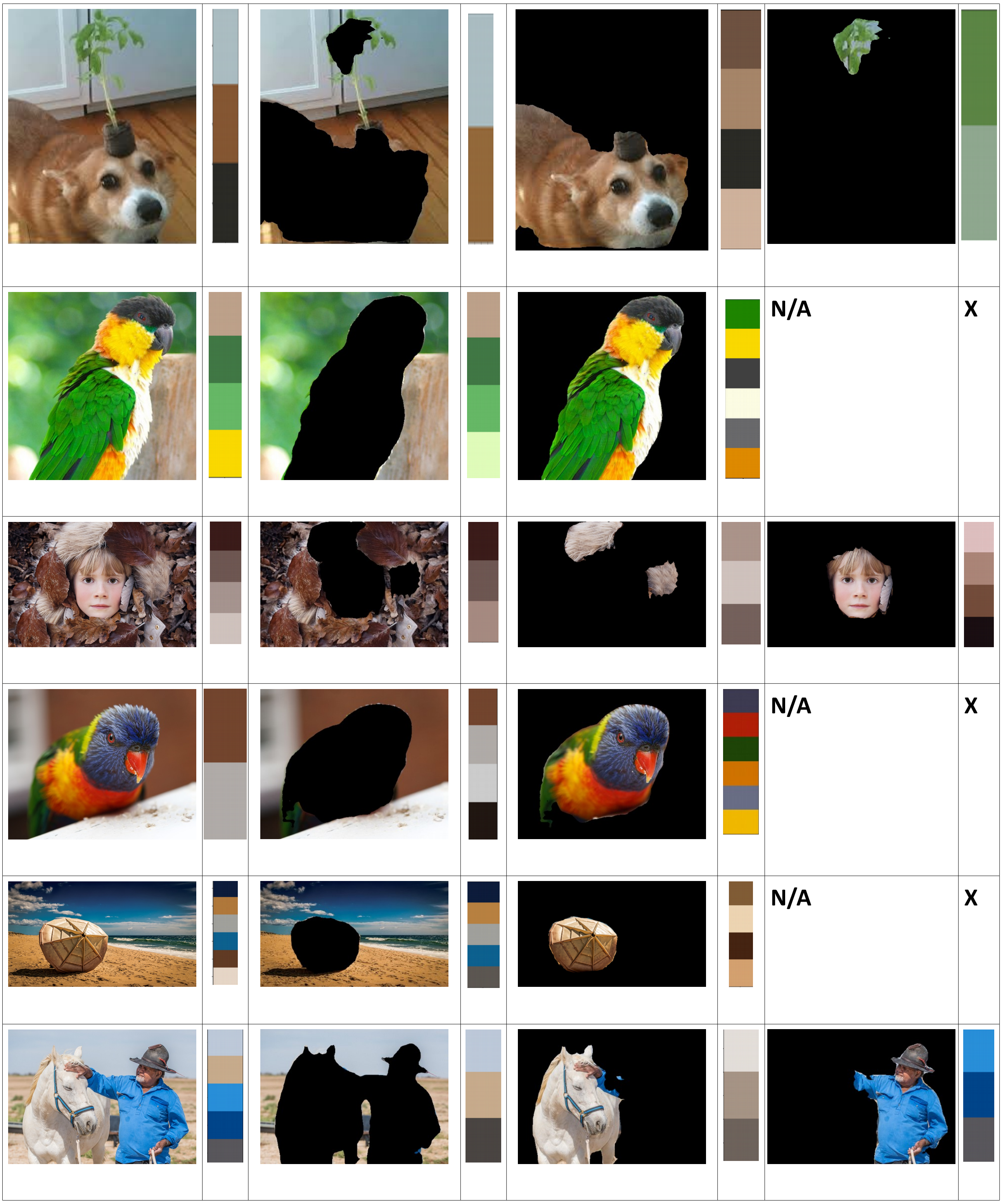
5					
6					**N/A**
7					**N/A**
8					
9					**N/A**
10					**N/A**
11					**N/A**

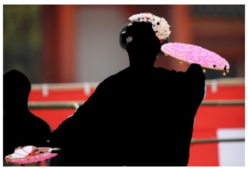
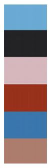
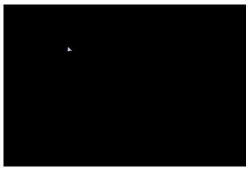
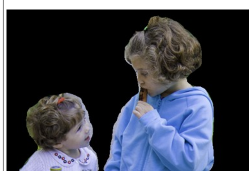
12					**N/A**
13					
14					

4.2 Palette Estimation

Table 4.2: Table of Image Segment Palettes

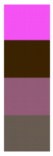
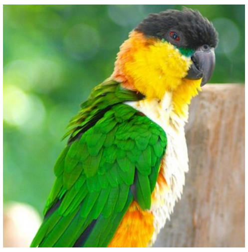
Original Image	P0	Segment 1	P1	Segment 2	P2	Segment 3	P3
						**N/A**	**X**
						**N/A**	**X**

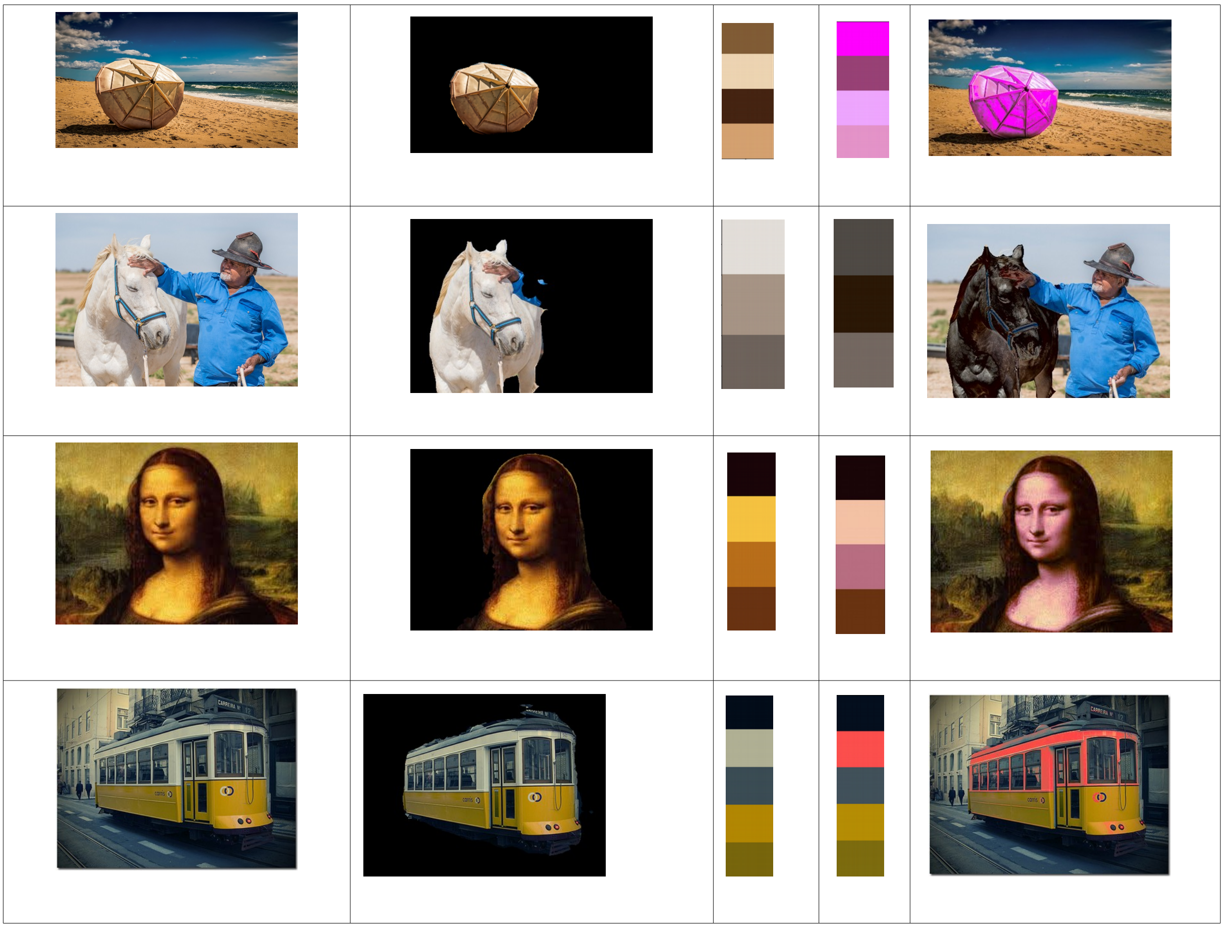

N/A
X
N/A
X
N/A
X

						**N/A**	**X**
						**N/A**	**X**
						**N/A**	**X**
						**N/A**	**X**
							
							

4.3 Combined Image

Table 4.3: Table of Image Alterations

Original Image	Segment used	$P_{Original}$	P_{New}	New Image
				
				
				
				
				



These results, are promising, showing that colour manipulation can be performed accurately with very little user input. That said, as can be seen in the images seen from the results, these methods are not perfect. The semantic segmentation can miss regions around the edges of objects, or find regions beyond the edges of objects.

These errors can give objects something looking like an aura surrounding objects, or slices of objects which do not get recoloured correctly. Clearer examples of these effects are shown in the portions of images shown in figures 4.1 and 4.2.

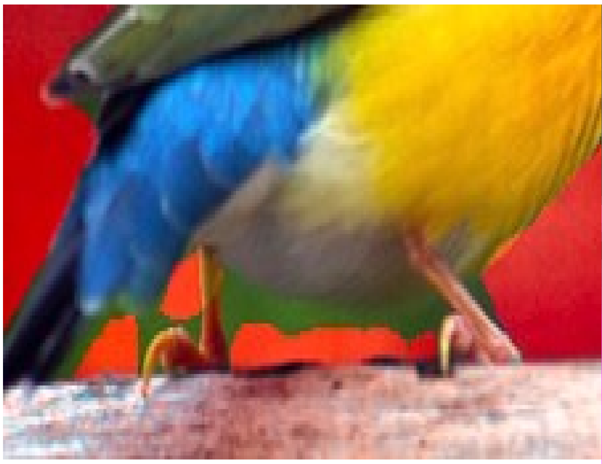


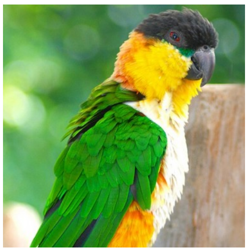
Fig 4.1: Region surrounding an object changed from Image 2

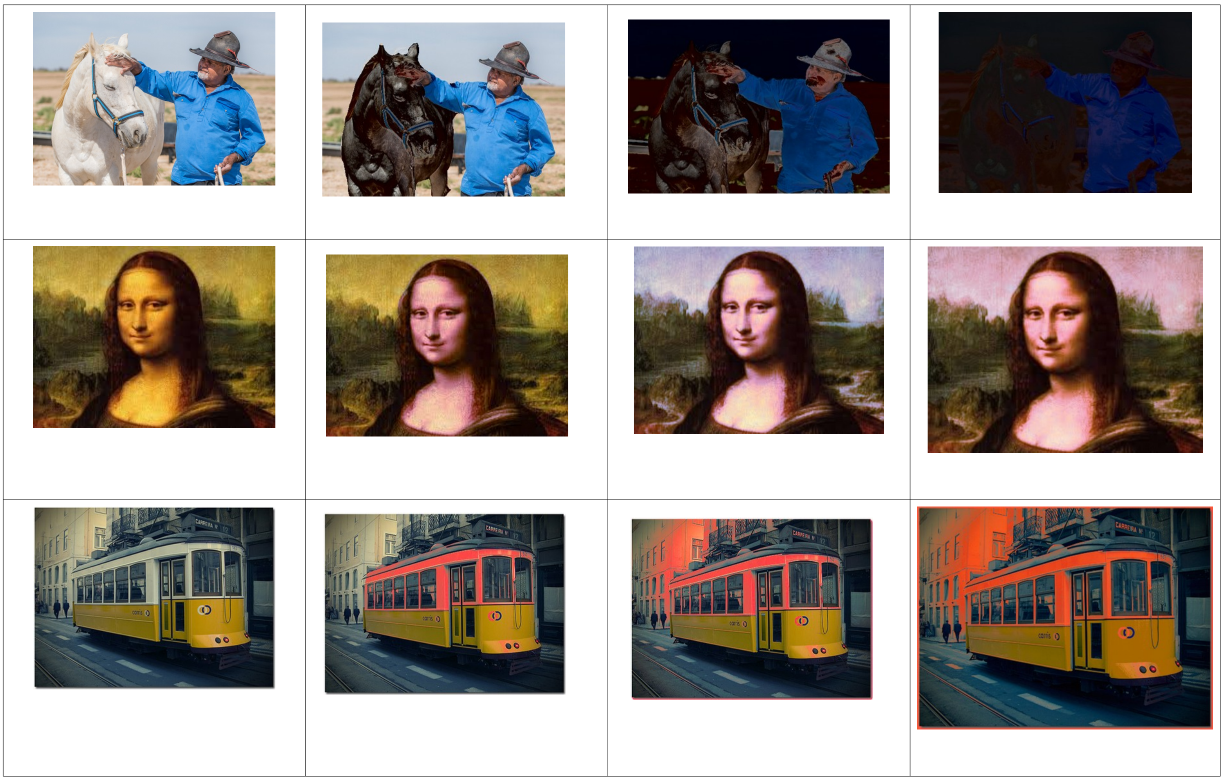
Fig 4.2: Slice of region not detected and unchanged from Image 1

4.4 Comparisons

While these results are not perfect, they do shown a marked improvement over simpler methods tried in the past. These results can be contrasted to results achieved when the same palette alterations are performed on the image as a whole, as would be typical using methods where the alterations are applied to the whole image. It should be noted that for the images shown, the palette used for Grogan et al.[1], all the initial and target palettes were identical to the palettes shown in the table prior. However for Chang et al.[14], estimations had to be made as the version of their method available online did not allow for specification of palette by exact values.

Table 4.4: Table of Comparisons between Palette-based Alteration Methods

Original Image	Proposed Method	Grogan et al.[1]	Chang et al.[14]
			
			
			
			
			
			



The primary areas which these comparisons show, are the areas in which the proposed method improves on previous methods. Particularly when control is sought for image manipulation, with regards to which region the manipulation is to take place in, with drastic, blatant changes being much easier to perform without negatively effecting the original image composition.

Chapter 5:

Conclusions

5.1 Design Applications

This approach is promising for the simplification of image manipulation software. It provides a simpler way to modify images relatively consistently, and even experts of image manipulation tools would find some method of semantic segmentation and object detection useful to speed up selection of image regions for alteration.

The palette estimation algorithm is a useful method for dynamic palette estimation. The values of the palette entries are typically representative of the perception of the image palette. Exceptions occur, such as the palette estimation for the overall image of Image 6, but these cases are few and far between, and these would likely be fixed by better selection of the area and threshold values used to define the regions selected.

Even when limited to the approach of palette based image manipulation, the results show that the range of good results that can be produced are much greater than without segmentation. Similar approaches may be suitable for some ranges of alterations, but cases where surroundings are of similar colours to the area which is to be altered can cause for surrounding colours to be altered along with the intended object.

Examples of this sort of effect are highlighted in the images taken from image 7 below, but can be seen across the images in the previous chapter.

Fig 5.1: Proposed method against non segmented method (magenta sand without segmentation)

The approach detailed in this report helps alleviate these issues with a degree of success, when accounting for small imperfections around the extremities of objects in an image. We can reliably conclude that the alterations are ideal for ~89% of pixels based on the accuracy of Deeplabv3.

5.2 Design Limitations

This approach has some limitations, which have been detailed in the results section of this report. As well as issues relating to segmentation regions themselves, there are also issue relating to internal separation of segments. For example, multiple objects of the same type in an image are assigned the same object value, and as such this approach does not separate individual objects from the scene. Also, while improvements have been made on methods of image manipulation that do not use semantic segmentation, the methods are still utilized by this approach, and as such many of the limitations are present in this approach as well. Examples of this are that when using the method of Chang et al.[14] when one palette entry is altered, the surrounding palette entries alter as well to adjust the image gradients, which may not be desired if the aim is to maintain certain areas of the image. This along causes image regions to get desaturated if multiple drastic changes take place.

5.3 Future Work

5.3.1 Soften segmentation region edges

While the approach is an improvement on global palette alteration, more work can be done involving the regions used for the manipulation of the image. The binary nature of these edge regions creates artefacts more often than desired, and a soft edge to these regions may help in amending this issue. These regions could allow for partial manipulations of palette around the edge of the segmented region of the image, such as in the example shown below:

Fig 5.2: Original segmentation mask

Fig 5.3: New segmentation mask

These fuzzy regions may account for errors in semantic segmentation, as well as removing small imperfections in the regions, such as tiny gaps. This may also help to remove regions of segmentation which belong to objects which should be part of the background, such as segment 2 of image 14, which, while correctly recognized as being pixels belonging to a chair, was useless due to it’s relative size to the rest of the image.

5.3.2 Analyze different colour spaces

The novel method for palette estimation is still not perfect, and more work could be done to analyze the performance of this type of algorithm in different colour spaces and with different shapes of ε. The current model uses a spherical region for the collection of data points in RGB space. LAB space in particular may also be worth exploring in this regard. As lightness is reduced to one dimension, it allows for oval regions of space to be included such that frequency of colour has slightly more weight than the brightness. This would make it so that similar colours could more accurately contribute to the same palette entry. Many regions of images when expressed in RGB space are seen as straight lines coming from the origin, allowing for multiple palette entries to be formed of the same clusters of colour when these gradient lines extend far enough through RGB space.

Along with this, much more work could be done with finding the ideal values for ε and θ for different applications. One area in particular which could be examined is altering the method to consider the energy of the colours in question, such that brighter and bolder colours need fewer pixels to reach the θ cut off.

5.3.3 Improvements to Semantic segmentation datasets

As semantic segmentation methods improve so too will this method of image manipulation. As the datasets improve so too will the methods of segmentation. As seen in the PASCAL dataset, only 20 object classes exist. This leaves a circumstance where it is nearly random whether and image not part of the dataset will be correctly segmented. Examples are shown below of cases where incorrect labels are assigned to objects in the image, both successfully and unsuccessfully.

Fig 5.4: Incorrect labelling, applied successfully

Fig 5.5: Incorrect Labelling, applied unsuccessfully

This is a product of both the segmentation methods as well as deficiencies in the datasets. An arbitrarily large dataset would allow for more correctly identified objects in a scene, however as the number of possible labels increases it may result in more incorrect labels. It may also lead to subregions of correctly labelled objects to being classified as different objects (such as individual wheels being classified rather than a car being classified a single object).

5.4 Closing Statements

Many approaches achieve their goal of providing a more simplistic version of image manipulation that allows for more precise control. The results, while not entirely objective, demonstrate this well, in spite of their shortcomings. From this we can conclude that the approach this method proposes is, if nothing else, a viable one. Improvements to the individual components of the approach will propagate into this approach, and as such the approach will become much more viable in the near future. As such it would not be surprising to see semantic segmentation or similar technologies represented in popular image manipulation software in the near future.

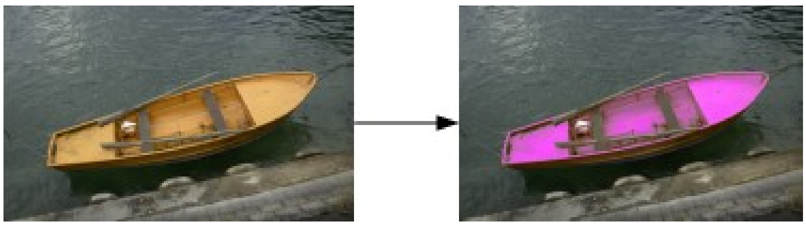


Fig 5.6: Closing Image of Proposed method

Appendices

A.1 Appendix of Figures

Figure ID	Figure Title	Image Source	Page Number
2.1	Segmentation map	http://host.robots.ox.ac.uk/pascal/VOC/voc2012/index.html	12
2.2	*Flow of operation in DeepLab CNN*	https://arxiv.org/pdf/1802.02611.pdf	12
2.3	Impact of Conditional Random Fields for semantic segmentation	https://arxiv.org/pdf/1412.7062.pdf	13
2.4	Image going through colour quantization (256, 32,16 then 8 colours)	Original work	14
2.5	Palette Estimation using median cut	Original work using "Color Thief"[10]	15
2.6	K-means clustering	https://en.wikipedia.org/wiki/File:KMeans-density-data.svg	16
2.7	DBSCAN Clustering	https://en.wikipedia.org/wiki/File:DBSCAN-density-data.svg	17
2.8	OPTICS Clustering	https://en.wikipedia.org/wiki/File:OPTICS-Gaussian-data.svg	18
3.1	Segmentation mask applied to separate image by object value	Original work	22
3.2	Randomly selected values in a 10 bin histogram	Original work	24
3.3	Local histogram to find optimal point value	Original work	24

A.2 Appendix of Tables

Bibliography

[1] Mairéad Grogan, Rozenn Dahyot, Aljosa Smolic
"User Interaction for Image Recolouring using L2"
Trinity College Dublin (2017)
https://www.scss.tcd.ie/~groganma/publications/CVMP17.pdf

[2] Touqeer Ahmad, Pavel Campr, Martin Cadık, George Bebis
"Comparison of Semantic Segmentation Approaches for Horizon/Sky Line Detection"
University of West Bohemia (2017)
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7966418&tag=1

[3] Mark Everingham (University of Leeds), Luc van Gool (ETHZ, Zurich),
Chris Williams (University of Edinburgh), John Winn (Microsoft Research Cambridge),
Andrew Zisserman (University of Oxford)
PASCAL Visual Object Classes Dataset (2005-2014)
http://host.robots.ox.ac.uk/pascal/VOC/

[4] Mark Everingham et al.
"The PASCAL Visual Object Classes Challenge: A Retrospective"
Internation Journal of Computer Vision (2013)
http://host.robots.ox.ac.uk/pascal/VOC/pubs/everingham15.pdf

[5] Liang-Chieh Chen, Yukun Zhu, George Papandreou, Florian Schroff, Hartwig Adam
"Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation"
Google Inc. (2018)
https://arxiv.org/pdf/1802.02611.pdf

[6] Mark Everingham et al.
PASCAL VOC Challenge performance evaluation and download server
http://host.robots.ox.ac.uk:8080/leaderboard/main_bootstrap.php

[7] Liang-Chieh Chen, George Papandreou, Iasonas Kokkinos, Kevin Murphy, Alan L. Yuille
"DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs "
Google Inc. (2017)
https://arxiv.org/pdf/1606.00915.pdf

[8] Paul S. Heckbert
"Color Image Quantization For Frame Buffer Display"
Massachusetts Institute of Technology (1980)
https://www.cs.cmu.edu/~ph/ciq_thesis

[9] Dan S. Bloomberg
"Color quantization using modified median cut"
Leptonica
http://www.leptonica.com/papers/mediancut.pdf

[10] Lokesh Dhakar
"Color Thief" (Python script used to perform median cut palette estimation)
https://github.com/lokesh/color-thief/

[11] J. MacQueen
"Some Methods For Classification and Analysis of Multivariate Observations"
University of California (1967)
https://projecteuclid.org/download/pdf_1/euclid.bsmsp/1200512992

[12] Martin Ester, Hans-Peter Kriegel, Jiirg Sander, Xiaowei Xu
"A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise"
University of Munich (1996)
http://www.aaai.org/Papers/KDD/1996/KDD96-037.pdf

[13] Mihael Ankerst, Markus M. Breunig, Hans-Peter Kriegel, Jörg Sander
"OPTICS: Ordering Points To Identify the Clustering Structure"
University of Munich (1999)
http://suraj.lums.edu.pk/~cs536a04/handouts/OPTICS.pdf

[14] Huiwen Chang, Ohad Fried, Yiming Liu, Stephen DiVerdi, Adam Finkelstein
"Palette-based Photo Recoloring"
Princeton University (2015)
http://gfx.cs.princeton.edu/pubs/Chang_2015_PPR/chang2015-palette_small.pdf

[15] Qing Zhang, Chunxia Xiao, Hanqiu Sun, Feng Tang
"Palette-Based Image Recoloring Using Color Decomposition Optimization"
Wuhan University (2016)
https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=7859399&tag=1

[16] YAGIZ AKSOY, TUNC OZAN AYDIN, ALJOSA SMOLIC, MARC POLLEFEYS
"Unmixing-Based Soft Color Segmentation for Image Manipulation"
ETH Zürich and Disney Research Zürich (2017)
http://zurich.disneyresearch.com/~taydin/_resources/publication/scs.pdf

Note: All links above have been accessed and found available on 16th of May 2018