COVENTRY
UNIVERSITY

Faculty of Engineering, Environment and Computing

School of Computing Electronics and Mathematics

Master of Science Thesis

Understanding and Classifying Cloud Structures In Satellite Images Using Supervised Deep Learning Techniques

Author : Abas Mohammedidris (8931834)
Supervisor: Dr. Aamer Saleem

Submitted in partial fulfilment of the requirements for the Degree of Master of Science in Data Science and Computational Intelligence

Declaration of Originality

I declare that this project is all my own work and has not been copied in part or in whole from any other source except where duly acknowledged. As such, all use of previously published work (from books, journals, magazines, internet etc.) has been acknowledged by citation within the main report to an item in the References or Bibliography lists. I also agree that an electronic copy of this project may be stored and used for the purposes of plagiarism prevention and detection.

Statement of copyright

I acknowledge that the copyright of this project report, and any product developed as part of the project, belong to Coventry University. Support, including funding, is available to commercialise products and services developed by staff and students. Any revenue that is generated is split with the inventor/s of the product or service. For further information please see www.coventry.ac.uk/ipr or contact ipr@coventry.ac.uk.

Statement of ethical engagement

I declare that a proposal for this project has been submitted to the Coventry University ethics monitoring website (https://ethics.coventry.ac.uk/) and that the application number is listed below (Note: Projects without an ethical application number will be rejected for marking)

Signed: Abas Mohammedidris Date:13/08/2021

Please complete all fields.

First Name:	Abas
Last Name:	Mohammedidris
Student ID number	8931834
Ethics Application Number	P122776
1st Supervisor Name	Dr Aamer Saleem
2nd Supervisor Name	Dr Diana Nkantah

Abstract

The advancement of technology based on artificial intelligence has witnessed a huge growth in recent years, this technology is not just improving our daily lives, it also has a crucial role in many roles such as saving lives and millions of dollars. One of the main concerns that is affecting our planet, is climate change, which directly affects our lives. Climate change is one of the main topics discussed by the forefront decision makers in recent decades. Clouds has a crucial role in climate and that is by controlling the amount of our plant's energy radiated back to space and solar energy that reaches the surface of our planet. Therefore, it is very important to understand the physical structure of clouds and how it is patterns are formed. This will highly contribute to improving climate model generation, which will result in an efficient climate change prediction as well as weather update. For thus reasons, this thesis implemented five different deep learning segmentation models for classifying cloud into four main classes: Sugar, Flower .Fish and Gravel. All the segmentation models were evaluated using dice coefficient and Jaccard Index (IoU).This study showed that using U-Net with good encoder, it is possible to have good cloud structures classification. Two of the models used in this project used scaled neural network EfficeintNet. However, the encoder can be replaced by another network.

Keywords: *cloud structures, satellite Image Segmentation, deep learning,U-Net, Dice Coefficient ,IoU.*

Table of Contents

1.1 Table of figures

1.2 Table of Tables

1.3 List of Symbols and Abbreviations

DL -Deep Learning

ANN -Artificial Neural Network

AI -Artificial Intelligence

ISSI - International Space Science Institute

SVM -Support Vector Machine

IoU -intersection over union

DC -Dice coefficient

ResNet -Residual Network

ResNeXt -Aggregated Residual Network

Acknowledgements

I would like to use this page to express my gratitude to all who supported throughout this journey.

I would like to express my sincere and deepest gratitude to my supervisor Dr Aamer Saleem who always inspired me and provided me with unlimited support and guidance for the completion of this thesis. I would like to thank alma mater Coventry University, especially my lecturers for their unlimited support, guidance, and encouragement, without them, all this work wouldn’t be possible.

Many thanks to my family, friends and everybody who supported throughout my journey.

Greatest thanks go to his Almighty Allah for providing me with the strength ,patience and everything I have in my life.

I would like to dedicate this thesis to my family for their unlimited support and love.

2 Introduction and Background

2.1 Introduction

One of the most interesting features of our planet as observed from outer space, is the distribution of clouds. Clouds are one of the most natural things that we encounter in our lives in a daily basis. Even though, clouds float above us, we hardly give the presence of clouds an attention and have a second thought and think about its influence in our lives, despite of the great impact of clouds in our planet's energy balance, weather, and climate. Therefore, clouds is an important field of study, for the purpose of understanding the Earth-Sun system (N.d, 2021) [1].Hence, having a better understanding of the clouds structures is very important for climatologist to get deeper insight into the weather of our planet. The reflection of energy without being absorbed is measured in "Albedo", for instance, most energy is reflected by white surface, which indicates high albedo, in the other hand dark surfaces reflects low albedo, which absorbs most energy. The warming of the climate is indicted by the earth's albedo, which is 0.3 albedo (Twomey, 1974) [2].

Understanding and interpreting cloud's structures will give a better insight into the abuse of our planet, in terms of climate and the risks associated with climate change. To understand cloud's structure, a broader picture of the atmosphere is needed, and this can be achieved by using satellite images, which give information of the earth current situation.

Climate change is affected by clouds in various way , which depends on the cloud's types such as, if the cloud types is thinner or thicker , higher, or lower on altitude and more or less abundant. Cloud has several types in terms of structure and one of the top-level and most abundant cloud types are Cirrus(Dowling & Radke, 1990) [3]. These types of clouds, which is feathery in structure, and it can compose of snow and it contains of long thin streamer, which is also known as mare's tails.

The cirrus cloud structure can indicate several weather conditions depending on its formation. For instance, if it's scattered in a fewer manner, it is an indication of good weather, however, if it increases gradually is a sign of humid storm and this could result into storm.

Another form of cloud is Cirrostratus clouds, which has form of thin scattered lines and its spread all over the sky and it gives the air white appearance. These clouds is a sign of rain approaching and though with these types of clouds the sky and sun can be seen clearly.

The Cirrostratus is a type of high cloud and it's a group of clouds white streaks. However, during summertime, they could indicate the approach of hurricane (McLean, 1957) [4]. There are many types of clouds forms that indicates the weather condition and these indications cloud be used to determine if any natural disaster is approaching and can be avoided beforehand in a low-risk manner. Therefore, the patterns formation of cloud and classifying their structures in advance allows to avoid high-risk activities such as aircrafts flights disaster, which carries thousands of passengers around the globe. Cargoes ships also heavily rely on the weather, for instance, if storms are encountered in the sea weather, this could cause shipments delay which results in the loss of assets worth millions of dollars. Beforehand prediction of the weather conditions could save the loss of lives as well as assets. Interpreting cloud structures in satellite images requires a well-trained expertise in meteorology. Although it is not clear and always possible to read, therefore an automated system with high intelligence is a good alternative. Hence, in this study five different models are developed and evaluated to classify cloud into four classes using deep learning segmentation techniques.

2.2 Background to the Project

In recent years the use of satellite images has increased rapidly for the purpose of providing high level detailed information in different applications, such as geographical information, urban monitoring, vegetation detection, weather prediction and too many other applications. Some of the key factors, which highly contributed to the increase usage of satellite images, are the improvements in the spatial and spectral resolution of satellite images. However, the improvement of satellite images in terms of spatial and spectral resulted in high complexity in satellite images processing for the purpose of classification, object detection or land cover mapping.

The resulted complexity due to the high spatial and spectral resolution can be described by the name of high intra-class spectral variation and this description is due to the image properties such as shadow and occlusion, details, and the increased date volume of the image in the processing stage, which is the result of the high radiometric, spectral, and spatial resolutions (Shu, 2014) [6] . As result of these challenges, the need of developing a new technique that can improve the process of satellite images has accelerated (Cai & Liu, 2013) [7] .

As a result of various attempts by researchers, many techniques was proposed and applied in satellite imagery, and the most important methods that were used are machine learning and deep learning techniques, and the use of the mentioned techniques has raised immensely with the increase power of classifier to discriminate objects within satellite imagery.

The problem under consideration on this research is, understanding and classifying clouds in satellite images .Clouds play a vital role in climate and directly contributes to climate and this is by directly controlling the solar energy that reaches earth's surface as well as the amount of the planet's energy radiated back to space. Therefore, the energy radiated back to space or the solar energy can affect climate in a positive and negative way, such as, if the amount of energy trapped inside the planet is very high, will result in a warmer atmosphere that will result in ice meltdown which increases the sea level. In the other hand, the less energy trapped in earth results in colder

atmosphere, hence understating cloud's structures gives intuitive insight of earth's weather. Therefore, it is very important for climatologist (Turner et al., 2007) [9] .
Regardless of image resource , whether , the image is collected from satellite or produced by a model output, more often it has sufficient information for scientist for the purpose of identifying features of interest , this process of feature identifying by scientist can be done for many images , in the basis of patterns across the images.
The human ability to identify features are successful in a situation where, there's features. However, the difficulty is encountered, where some patterns are complex and difficult to describe objectively, in a situation, where the purpose is explicit method for pattern identification, this results in a frustration of developing the pattern identifying methods. In this situation machine learning comes into play and more specifically deep learning, which already demonstrated the ability to outperform the human ability in identifying patterns (Muhlbauer et al., 2014) [10].However , the use of deep learning application requires sufficient dataset for training and evaluation ,where limitations are encountered specially in cloud dataset.
Therefore, the limitation of cloud dataset led to various methods in combining and labelling cloud dataset, one of them being the dataset considered for this study, which was published on Kaggle competition by [11], under the name of understanding and classifying cloud structures. The most common feature of satellite imagery is mesoscales patterning of shallow cumulus, however , the organisation on this scale is heavily ignored ,when it comes to modelling studies of clouds and climate (Rieck et al., 2012) [12] . The fact of mesoscale patterning of satellite imagery led the International Space Science Institute (ISSI) team to identify mesoscale clouds patterns that are frequent in the northern Atlantic. The ISSI team named the patterns Sugar, Fish, Flower and Gravel as it is shown in Figure 2.1 . The ISSI names previously were described in the literature, below show the relation of ISSI patterns, with previously identified patterns.
"Sugar" describes in areas of very fine cumulus in widespread manner, this class of clouds exhibits the mesoscale organisation, because the overall field does not have much of cloud-free region. More often are embedded within very large flow that gives sugar some structure. In the event of strong flow, this class or pattern form thin feathers, which was previously described as dendritic clouds.
"Flower" the areas described by flower are the clouds that have isotropic structures , which have a range of 50km to 200 km in diameter , which consist of similar cloud-free region in between, the flower are less densely packed , and it overlaps to some degree with previously described patterns stratocumulus and closed-cell MCC (Norris, 1998) [13] .

"Fish" the fish feature or patterns are elongated ,which sometime spin up to 1000 km , in most cases in longitudinal manner, in previous literature this pattern was called action-from clouds by (Garay et al., 2004) [14] .

"Gravel" this pattern name describes the fields that have granular features whose are marked by rings. The normal scale of these rings is approximately 20 km, it is highly suspected that, this pattern is driven by cold pools that are caused by raining cumulus.

Previously, researchers attempted identifying clouds patterns in various ways, for instance, scientist detecting cloud patterns, which could be successful for small number of images, but that's not always the case. These mentioned challenges makes deep learning of high importance, for the purpose of segmenting and classifying clouds in satellite images. Therefore, this study will develop, implement, and evaluate the latest state-of-art deep learning techniques proposed by researcher for image segmentation and classification.

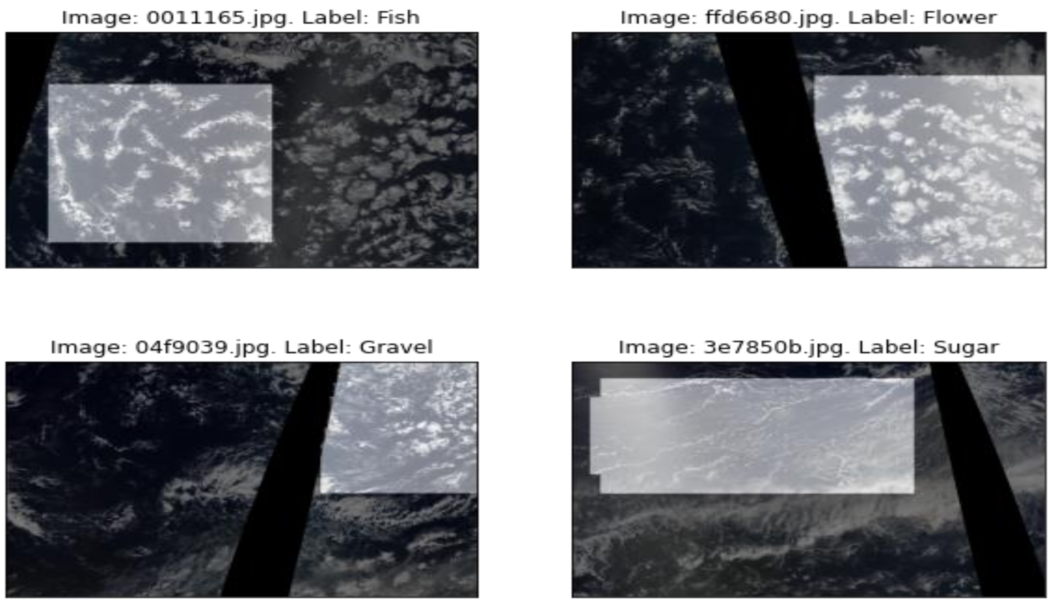


Figure 2.1 four cloud types

2.3 Project Aims and Objectives

This project aims to use segmentation models in classification task of satellite images and the main purpose is classifying clouds structure into four main classes shown in Figure 2.1 and the classes names are Sugar, Flower, Gravel, and Fish. The objectives of this study is to develop five different segmentation models and apply them , in classification task as well as evaluating the models on the underlying dataset and determine the most appropriate model for the application of cloud classification using supervised deep learning. After evaluating the models and conducting comparison, then conclusion will be drawn based on the attained results and recommend future work can be done to improve the results. The dataset set image size is large and requires a powerful computing and this study aims to adapt and explore possible solutions.

2.3.1 User Requirements

Cloud classification in satellite images is crucial in many fields and applications. The main industries that will benefit from this study are cargo shipment via sea, aviation industry and more importantly climatologist working on research such as climate change. The cargo ships are highly dependent in the sea weather and they are subjected to face weather uncertainty such as storms and this could result in the loss of millions of dollars due to shipment delay or could result in a worse outcome such as lives loss. Another industry that highly rely on the weather is the aviation industry, aircrafts carry millions of passengers across the globe at a specific time interval and flying through cloud , is the same driving in fog, which cause lack of visibility, which makes pilots loss the ability to see what is ahead. To avoid the mentioned risks and save lives as well as millions of dollars, in that case the main requirement is to predict clouds structure with high accuracy beforehand, when it is formed. This will assist the mentioned industries.

2.3.2 Analyse and model the requirements

This study proposes and implements segmentation models for classifying clouds structure in satellite images into four main classes. However, classifying clouds structure in satellite is challenging due to the high similarity in clouds patterns. The similarity between clouds patterns makes the task of classifying it into different classes using traditional classification technique very challenging and unreliable .Possible solution is by considering all the factors which include the requirement of users and the challenges that will be encountered and provide an appropriate solution.

2.3.3 Possible solutions

To achieve the user requirement, this study uses fully connected convolutional neural network for semantic segmentation and the reason behind choosing semantic segmentation models is, these models classify each pixel belonging to specific class or belonging. Five different models are developed in this study. Each model has its merit, all the models are discussed in great details in section 4.2. To understand each model and its suitability for this project, intensive literature review was conducted on previous research, which helped in narrowing the choice of the models. There are various techniques and models can be used on this project; however, the chosen model have the best performance in competitions such as google AI competitions and large images dataset such as ImageNet.

2.4 Overview of This Report

- This project report starts with an abstract, which gives overall insight to the project, followed by table of contents, figures, tables and list of abbreviations.
- The second section of this report is the Introduction and Background , which introduces the project and presents a background to the project. It also presents the project aims and objectives, user requirements, problem statement and possible solutions.
- The third section is Literature Review and Theoretical Approach, which presents related works and gives general overview of deep learning and theoretical approaches can be used in this project.
- The fourth section of this report is Methodology , this section presents data description and data pre-preparation as well as technical approaches, which introduces the models used in this project and give overall understanding about the models.
- The fifth section is Design , the design process of the project and design summary decision to complete the project and meet the requirement.
- This sixth section is Implementation ,in this section, the implementation process of the project is presented.
- Section seven is Results and Evaluation ,the section is about the evaluation metrics used to evaluate the model's performance as well as results attained to assess model performance.
- Section eight is Project Management ,which explains the project management to achieve the aims of the project, which is crucial to the project quality, it also presents risks associated with the provided possible solution the avoid the identified risks. Furthermore, it presents social and ethical consideration.
- Section nine is Critical Appraisal,which gives critical appraisal about the project.
- Section ten Conclusion ,which concludes the project , discuss the achievement of the project, and recommend future works.
- The final section is Student Reflections ,which gives overall experience gained from the project, such as challenges encountered, and they were solved as well as the transformable knowledge gained from the project.

3 Literature Review and Theoretical Approach

3.1 Literature Review

One of the most common and important meteorological phenomena are clouds, which covers approximately 66% of the global surface. Therefore, its analysis in terms of condition and coverage plays a vital role due to its importance in various applications. Localisation and cloud condition can be achieved simultaneously and accurately by using a ground-based observed clouds with high temporal and spatial resolution. Some of the ground-based devices for clouds measurements are radars, lidar and another method is satellites which orbits the surface of earth. Hence, accurate and efficient cloud structures segmentation has become a topic of interest for many researchers in recent years ,due to its importance to meteorologist and climatologist for further understating of weather and climate conditions .Therefore, many techniques and models have been proposed and developed for understanding, classifying and segmenting clouds images (Tapakis & Charalambides, 2013) [15] .

A deep CNN model for clouds images segmentation was proposed by (Xie et al., 2020) [16] , the model was named SegCloud , the proposed model consist of a symmetrical encoder-decoder architecture and it operates as follows , the encoder combines the low-level of the cloud features and forms cloud features with high-level and low-resolution, and the other network ,which is the decoder restores the high resolution features ,which was obtained and maps an output that has the same resolution as the input image. SoftMax classifier is the final stage of the network, which performs pixel-wise and gives an output of the segmentation result. The model achieved a promising result, when it comes to cloud segmentation on four different conditions, clear sky, partial cloud, overcast sky and average.

A colour-based cloud images segmentation technique was proposed and developed by (Dev et al., 2017) [17] , the method is entirely supervised segmentation, and it uses systematic analysis of components and spaces of different colours with least square regression. . Also, this method is based in learning, and this result in no requirement of its parameters to be defined manually. The technique was tested in Singapore whole sky dataset, which was created by the same authors.

(Yuan et al., 2017) [18] , introduced a Deep learning technique for efficient cloud detection for remote sensing application, which is based in multi-tasking, that is cloud segmentation and cloud edge detection for encouraging accurate detection, the authors also proposed an easy way of training their model and evaluated the model using IoU and other techniques and it was found that, the method returned a superior performance that outperform the previous state-of-the-art techniques.

A novel approach for cloud regions detection in images was proposed by (Shi et al., 2016) [19], the method is based in the state-of-the-art Convolutional Neural Network and it consist of six layers , four convolutional and two fully-connected layers , the model works as follow . Firstly, the image is clustered into super-pixels , which is, as a sub-region and this process is achieved by simple linear iterative cluster, which then generate the cloud map probability of the image and the region of the cloud is obtained in the final stage. The proposed method was successful in detecting thin and thick cloud with robustness and effectiveness.

(Azimi-Sadjadi & Zekavat, 2000) [20] , used support vector machine to classify areas with clouds and cloud-free area in satellite images, the SVM was used to classify 10 different classes, six classes, which consist of six different cloud types and four non-cloud classes, due to the limitation of SVM of only classifying two classes at the time of this study, the study designed hierarchy support vector machine, to extend its capability to classify ten different classes.

Furthermore, (Qing Zhang & Chunxia Xiao, 2014) [21] , presented an automatic algorithm for cloud detection in colour aerial images , the methods , firstly builds a significance map that shows and identifies the difference between regions with cloud and non-cloud and with the use of optimal thresholds , an output will be obtained for cloud detection whether its cloud class or non-cloud class. Then the process is furtherly proceeded to narrow the regions candidates with cloud and construct map for accurate cloud detection. The final stage of this method is detecting semi-transparent cloud pixels in the highlighted regions of the image, this method has good performance for detecting clouds in aerial colour images.

A texture-based artificial neural network classifier with only one single-channel, for the purpose of classifying clouds in spatial resolution data, was proposed by (Lee et al., 1990) [22] . This method achieved an outstanding result with overall accuracy of 93% in identifying clouds. However, this method did not consider the use of inferred channel and it classifies specific types of clouds with high accuracy.

A study by (Zhang et al., 2018) [23] , proposed a new model, under the name of Cloudnet ,for identifying cloud types , the model architecture is based on deep convolution neural , this study rely on three main things simultaneously, the shape features , texture and structure , the proposed model was tested and evaluated on dataset generated by the same study and achieves high efficiency in classifying 10 types of clouds and the main motive behind the development of this model was , to be deployed on ground-based clouds classification application.

A remote sensing network based in deep learning for cloud detection was introduced by (Jeppesen et al., 2019) [24] , the detection model has U-net architecture and it is designed to detect clouds in optical satellite images, the proposed model was trained and evaluated on dataset, that has very hard scenery to distinguish clouds, such as clouds in regions with snow and ice, the model showed a promising with RGB images and existing masking techniques. After evaluating the model with various evaluation methods, the study concluded that model is ready for production environment. A novel approach for cloud and shadow detection was introduced by (Zhai et al., 2018) [25], the technique is based on spectral indices and the aim of the algorithm proposed by this study is , to be used in a wide range in multi/hyperspectral in optical sensing applications with both spectral channels, visible and inferred respectively. The effectiveness of the algorithm was evaluated on 8 different hyperspectral sensors and in the demonstration achieved a promising result with accuracy of 93.13 in detecting cloud and shadow.

(Li et al., 2015) [26], introduced a new technique for cloud detection , the method is based on the state-of-the-art algorithm support vector machine ,the algorithm works as follow , the satellite images is divided into small blocks as well as the brightness characteristics sub-blocks is extracted for the purpose of initial detection , which is followed by the process of computing the co-occurrence average grey level and gradient second order matrix ,which is based on the texture features, for the basic SVM. The sub-block image with brightness characteristics is used as learning sample for the support vector machine classifier to initiate the final classification details, using the SVM. The experimentation results for this method gave an accuracy of 90% on cloud detection. (Wang et al., 2020) [27], performed two types of cropland segmentation using U-Net models , the two types of labels, whose are labels that consist of a single geotagged points and image-level labels , both labels types are commonly found in remote sensing imagery datasets ,which in most cases is considered as sources of “weak supervision”, the study also demonstrated that , U-Net that is trained on single pixel per image and U-Net trained on the weak label can outperforms logistic regression ,random forest and support vector machine ,whose are algorithms of pixel-level.

Additionally, study conducted by (Jin et al., 2014) [28] proposed a novel technique for classifying clouds in satellite imagery , a method that uses over-complete dictionary which is motivated by sparse representation , for the purpose of representing clouds sample sparsely, the process is followed by extracting the dictionary feature using sparse coding. To form subspace projection of cloud, the sparse coefficient matrix is set to specific cloud of the training samples ,which completes the sparse classifier design ,the classifier was tested against the state-of-the-art classifier SVM and it has better performance than SVM in identifying different conditions of land on satellite imagery such clear water ,clear land, and different types of clouds.

(Ahmed & Nuri Sabab, 2021) proposed an encode-decoder based deep learning for clouds classification , the models use EfficientNet as encoder and the original U-Net decoder and decoder, the final evaluation for the proposed models were evaluated and achieved for EfficientNetB0-B5 a mean accuracy of 0.6389 , 0.6423 ,0.6351 ,0.6311 , 0.6294 and 0.6255 respectively. The final model is called EfficientUNet, which consist of encoder and decoder in u-shaped architecture.

3.2 *Theoretical Approach*

3.2.1 Deep Learning

Artificial neural network is a computing system that is made up of several simple and highly interconnected processing elements .The ANN process information by their dynamic state response to external inputs and it is biologically inspired by the human brain. Despite of its existence for decades, recently caught the attention of researchers in the computer vision, machine learning and artificial intelligence communities. There are various types of artificial neural network such as single-layer neural network and an example of this type of NNs is the Hopfield network. Multiple feedforward neural network such as functional link, standard backpropagation, and product unit network. Temporal networks for example time-delay neural networks as well as Jordan and Elman simple recurrent neural network. Self-organising NN and an example of this network is learning vector quantizer and Kohonen self-organizing feature maps. Furthermore, the NNs which is the combination of feedforward and self-organising neural networks and an example of this NNs is the radial basis function network.

Now days, artificial neural network is used in various applications such as classification, and the purpose of it , to make prediction of class of an input vector and more applications of ANN such as pattern matching, pattern completion, optimisation, control, function approximation or time series modelling and date mining(N.d, 2018) [29] . In terms of structure, artificial neural networks can have feedforward neural, feedback networks and lateral networks. An example of feedforward networks is shown in Figure 3.1.

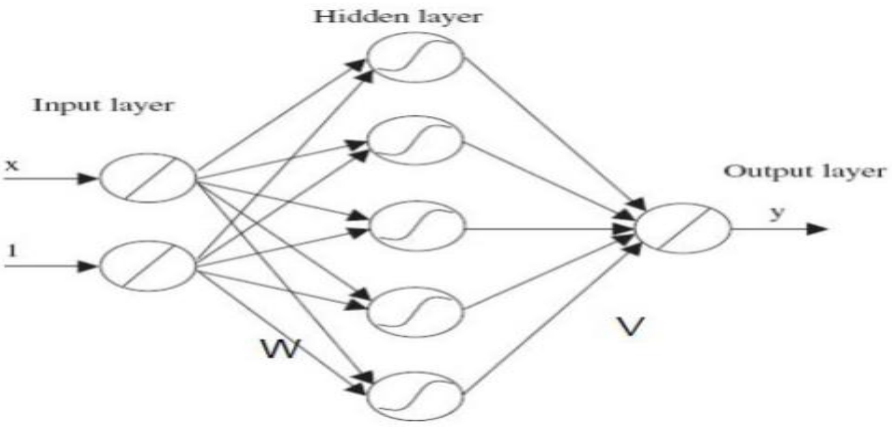


Figure 3.1 : feedforward Neural Network 51

The above network size is n × m × r, where W_{mn} is the input weight matrix and V_{mn} corresponds to the output weight vector, the network does not consist of feedback network.

3.2.2 Convolutional Neural Networks

One of the most well-known deep learning algorithms, which has gained attention in recent years for image processing is convolutional neural network (CNN), this algorithm has superiority over other deep learning neural network due to its ability to keep the image geometric shape such as (2D format) and specifically has the capability of maintaining the interconnection among pixels and keep the spatial information (Zhu et al., 2017) [30] .This section provides a short overview of convolutional neural network in terms of architecture and working process.

3.2.2.1 Architecture of Convolutional Neural Networks

By considering CNN is applied to an RGB image which is 3-dimensional matrix like the images under consideration for this study. Therefore, the CNN takes note of the image considering its dimensionality, which is in fact 3-dimensional as it is shown in Figure 3.2.Each layer of the convolutional neural network consists of filters or kernels with specific volume , which is also called neuron and its size is h × w × d where “d” and “h” corresponds to the spatial dimensional of the filter and “d” corresponds to the number of channels or features.

For instance, Grey scale image has 1 channel and RGB as mentioned 3 channels. Each one of the kernels is convolved and sliding across the entire image taking into consideration the input image volume and the sliding is with size H_i , W_i and D_i where H_i , W_i are the spatial dimensions and D_i the number of channels.

The term convolution refers to the summation of neurons in element-wise dot product principle in each kernel or filter with corresponding values of the input. The input image is considered as the first layer and the output of single layer convolution is 2-dimensional output of specific size and its determined by the parameters such as padding and stride, this procedure is illustrated in Figure 3.2.The stride of the filter s can be defined as the filter interval moving in each spatial dimension and the padding can be defined as p_h and p_w which corresponds to the pixels added to the outer edge of the input. Hence the stride can be considered as the means of subsampling. In most cases square matrix kernels or filters are used and an example of that $h = w = f$ and the corresponding output of such layer can be determined by equation 3.1 and 3.2 .

$$H_o = \frac{H_i - f + 2p}{s+1} \tag{3.1}$$

$$W_o = \frac{W_i - f + 2p}{s+1} \tag{3.2}$$

Where $D_0 = N$

It is very important that the size of the filters should not be the same in various layers.

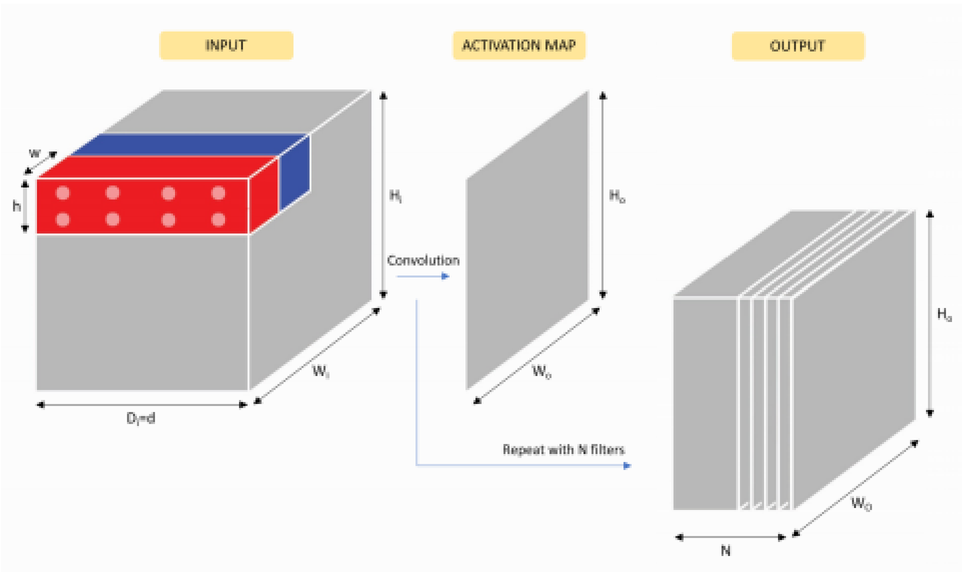


Figure 3.2: illustration of single layer CNN

Assuming the data under consideration is an RGB image with 3-dimensional matrix , the red and blue areas corresponds to the same filter in two positions with a size of h × w × d, this filter is convolved across the entire input volume by sliding it. The output volume is attained by using N filters , for explanation let's consider the filter has 2 × 2 and the striding parameter s =2 and from Figure 3.2 the RGB image input is D_i =d =3 .

3.2.3 Activation Functions (Non-linearity)

Activation functions map the input to the output in neural network, whereas the input values are attained by computing the weighted sum of the neurons input and additionally adding a bias if there's bias. Which means the activation or non-linearity function will decide if specific neuron within the network will fire or not for a given input by outputting a corresponding output.

After each learnable layer in convolutional neural network architecture, for example convolutional and fully connected layers, there's activation function used. The CNN can learn more complex things with non-linearity behaviour and manages to map the output non-linearly. One of the most important attributes of non-linearity functions, is differentiability and it is differentiable will enable error backpropagation for the purpose of training the model. Some of the most common activation function used with deep learning are reviewed in the following section.

3.2.3.1 Sigmoid function

The working process of the sigmoid function is as follows, it bind the output in specific range such as [0, 1] and that is achieved by taking real numbers as input. The sigmoid function has a shape of 'S' and it can be represented mathematically by equation (3.3).

$$f(x) = \frac{1}{1+e^{-x}} \tag{3.3}$$

3.2.3.2 Tanh

Normally, the tanh function is used as an activation function that bind the output within specified range of [-1,1] and it takes real numbers as input , the tanh function is represented mathematically by equation (3.4) .

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \tag{3.4}$$

3.2.3.3 Rectified Linear Unit (ReLU)

In convolutional neural network, ReLU is the most used activation function, and it is used for the purpose of converting all the inputs values into positive numbers, the main advantage of using this activation function is, it has less requirement of computation load than the other functions and the mathematical representation of the rectified linear unit function is represented by equation (3.5).

$$f(x) = \max\,(0, x) \tag{3.5}$$

3.2.3.4 Leaky Rectified Linear Unit (Leaky ReLU)

The difference between the Leak ReLU and ReLU is that the LeakyReLU does not ignore a negative input entirely instead down-scales the negative, mainly LeakyReLU is applied in ReLU dying conditions and it is represented mathematically by equation (3.6).

$$f(x) = \begin{cases} x, & if\ x > 0 \\ mx, & x \quad \leq 0 \end{cases} \tag{3.6}$$

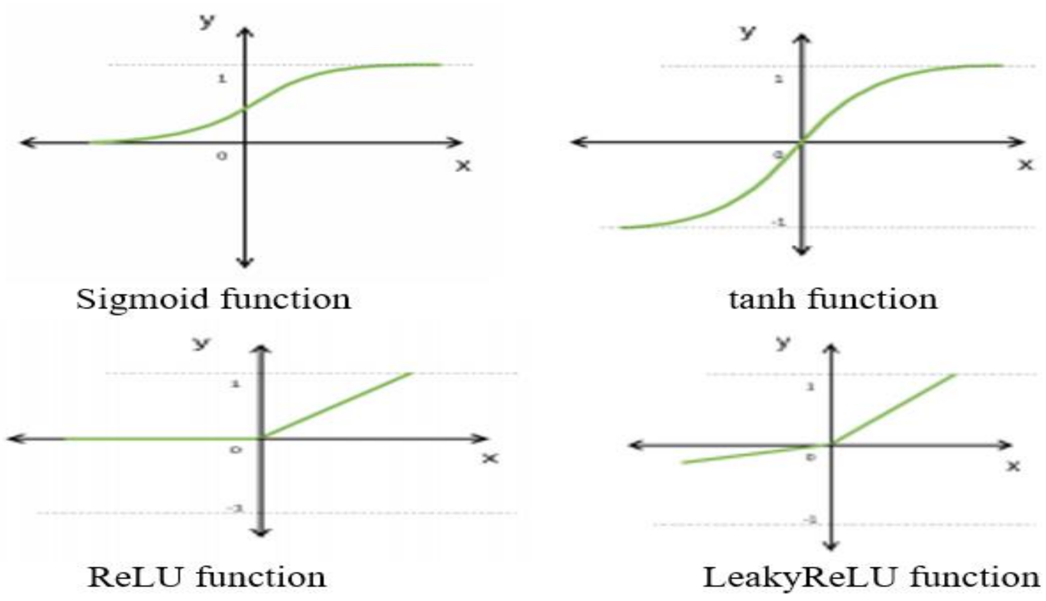


Figure 3.3: Activation functions

3.2.3.5 Pooling

Pooling layers are applied to sub-sample the feature maps, the purpose of using pooling layers is to take large size feature maps and diminish them into feature maps with smaller size, in the process of decreasing the features maps, and it always keeps features information with higher dominancy. Pooling is conducted by initially pooled region size as well as the stride of the operation, like the convolution operation. Various techniques of pooling are used depending on the problem on hand for example min pooling, average pooling, max pooling, gated pooling, tree pooling and so on. However, commonly used one, is max pooling, the only disadvantage of pooling is, in some cases decease the model overall performance. A basic operation of pooling is presented in Figure 3.4.

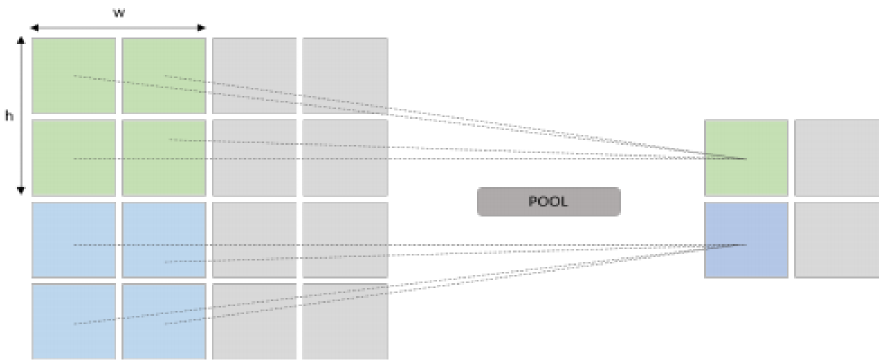


Figure 3.4: Illustration of pooling

3.2.4 Fully connected layer

Following the process of detecting the features with high level from the proceeding of convolution and pooling layers, normally the final stage is the attachment of fully connected layer at the end of the network. The neurons of the fully connected layer are connected to the entire input volume received from the convolution, pooling, or activation. This layer determines the features that corresponds to each class and that's achieved by taking all activations received into consideration. The neurons activations is determined by using equation (3.6)

$$y_i = W_i x_i + b_i \tag{3.7}$$

Where x_i denotes the input vector , W_i is the corresponding weight matrix and b_i is the bias vector .The disadvantage of fully connected layer is ,the neuron receive activations from all the input neurons , therefore , spatial information is lost , which is not desired in semantic segmentation. Hence, the fully connected layer can be as considered as 1x1 convolution, which will be applied to the entire input with full-connection mapping. The filters can be considered as having an equivalent spatial extent to the input layer (Muruganandham, 2016)[31].The illustration of fully connected layer converted to 1x1 convolutions as shown in Figure 3.5

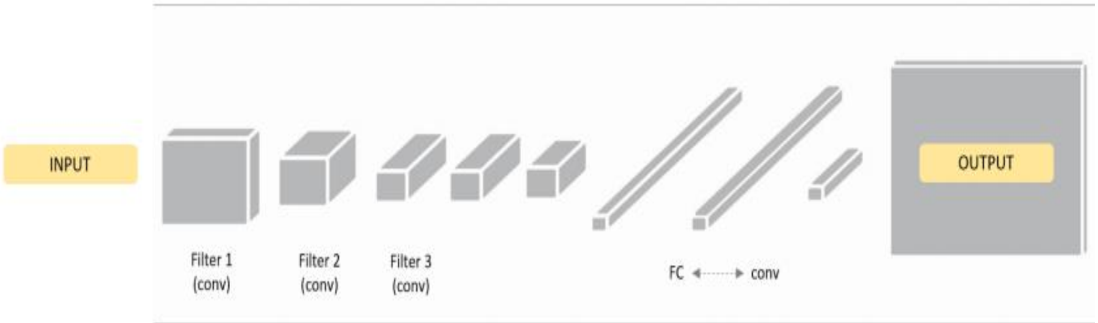


Figure 3.5: converted fully connected layer to 1x1 convolutions [31]

3.2.5 Regularisation

One of the main challenges with dep learning is to adapt properly the unseen input data, which is different to the training or the seen date, both data are drawn from the same distribution, and the ability of adapting the unseen data is called generalisation. CNN models are powerful enough, where in some cases fit itself on the training and fail to achieve similar performance on the unseen data. For instance, if the model performance is good on the training dataset but fails on the unseen data, this type of model is known as over-fitted and when the model did not learn enough then it is called under-fitted as well as, if a model performance is good on both training and testing then it's called just-fitted (Ghosh et al., 2019) [33] . An example of the three cases with respect to binary classification is shown in Figure 3.6 .

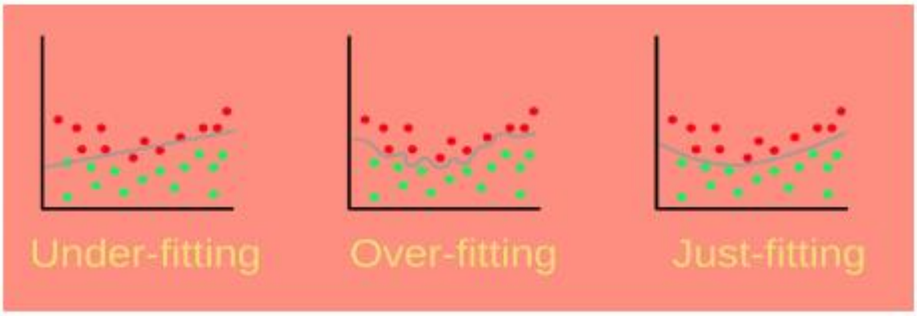


Figure 3.6: Illustration of under-fitting , over-fitting and jus-fitting [33]

3.2.5.1 Dropout

The most regularisation technique used to avoid over-fitting is dropout, which was proposed by (Srivastava et al., 2014) [32]. The technique working process is as follows, in each training epoch, neurons from the network is randomly dropped and this helps to distribute the feature selection equally in all the neurons and forces the model to learn different independent features. The dropped neurons will not take place on the training process on both backpropagation and forward propagation. For performing prediction, the full-scale network is used. The effect of dropout in the network during training process is demonstrated below in Figure 3.7.

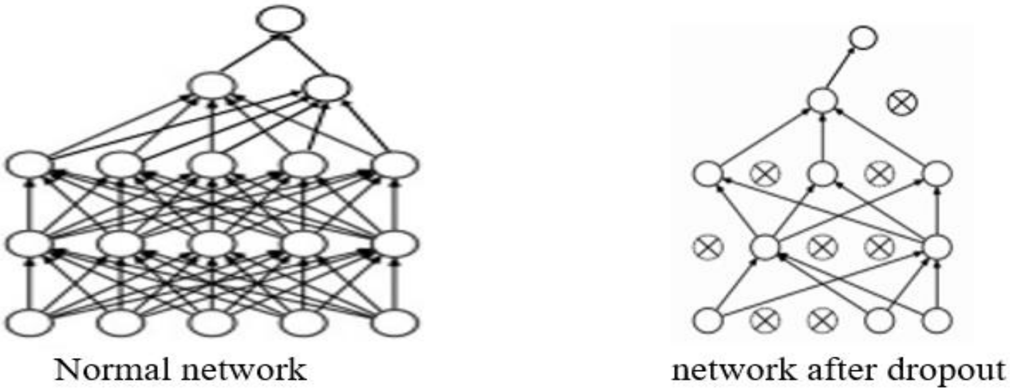


Figure 3.7: Illustration of dropout [32]

3.2.6 Semantic Segmentation

Throughout the years, the interpretation of visual information has been approached in various ways, however the aim remains the same, examining and investigating images to identify objects and evaluate its significance. The techniques to learn from visual information is classified into these main categories , image classification ,semantic and instance segmentation , object localisation and detection and others (Muruganandham, 2016)[31] .However this study is mainly focused on semantic segmentation for satellite images. Hence, this section will present a general introduction and review to semantic segmentation. To have a better understanding of semantic segmentation, let's briefly define image segmentation and types of image segmentation. As the data of interest is image and image is nothing except of the fact that is a collection of pixels. Therefore, the process of classifying each pixel in the image to certain class belonging is called image segmentation and it can be said as problem per pixel. There are two types of image segmentation.

- Semantic segmentation: which is the under consideration in this study, this technique classifies each pixel belonging to specific class or label , for instance , if an image consist of two dogs , this method will give the same label to both dogs within the image.
- Instance segmentation: This technique varies from semantic segmentation because it gives a unique label to every instance of specific object within the image.

For semantic segmentation purposes , an end-to-end model that composes of two linked parts was proposed by (Noh et al., 2015) [34], the model consist of two parts , the first part is VGG16 model architecture ,which was used as convolutional network ,which takes an input proposal such as bounding box which can be generated by model which detects objects. Then a vector of features is generated by processing and transforming the proposal using convolutional network. The second part is the decoder or the deconvolutional network, which takes the vector of features as its input and generates a map of pixel-wise probabilities to determine the belonging of each class.

Furthermore, it uses unpooling which targets the maximum activations and preserve the locations of the maps. It also uses deconvolution to expand the feature maps while preserving the dense. This network achieved a mIoU of 72.5 on segmentation challenge in 2012 PASCAL VOC challenge. The full network architecture is shown Figure 3.8. This technique was further expanded to create the U-Net architecture, which is discussed in section 4.2.2. The same approach is adapted to create all the models used on this study.

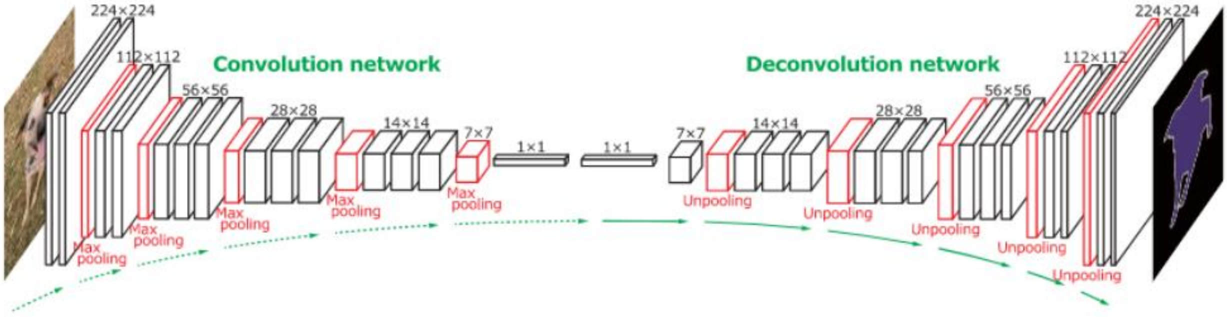


Figure 3.8: Two parts linked parts segmentation network proposed by (Noh et al., 2015) [34]

4 Methodology

To complete the project within the planned time frame and better performance and the for the completion of product, the agile methodology was used. The agile methodology breaks the whole project into small phases of work and this technique is also known as feature and produces what is called a Minimum Viable Product. By using iterative process, the Minimum Viable product is then gold plated. This method also collaborative between the stakeholder and take their inputs at every stage of the work and continues improvement [35].

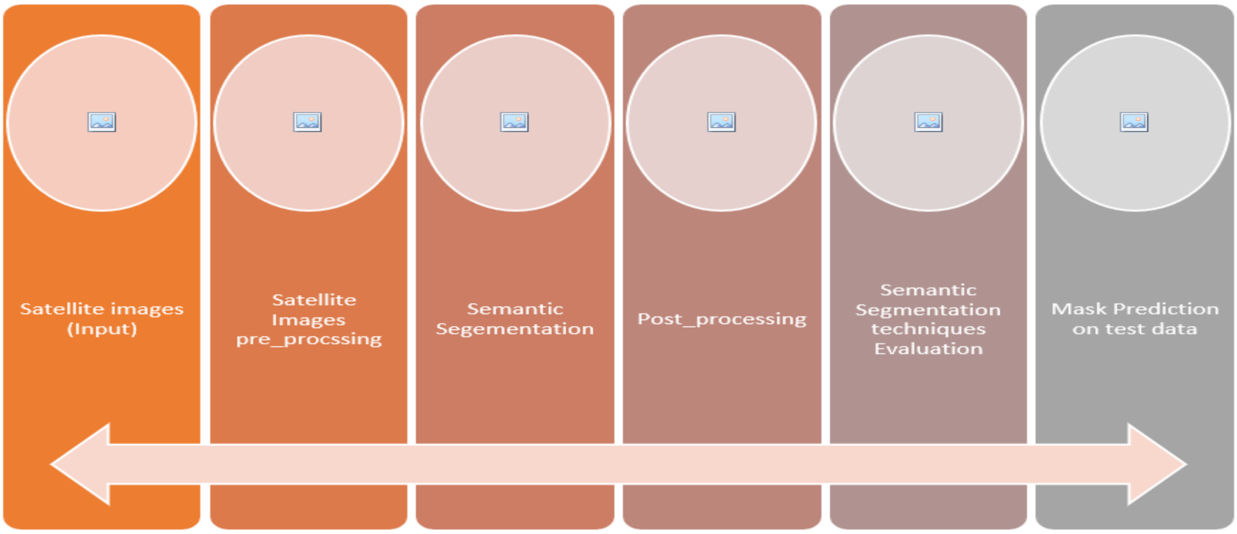


Figure 4.1: project feature maps

The first stage of the project is the pre-processing of the data which includes , date exploratory data analysis (EDA), data preparation , data cleansing , data augmentation , label decoding because the labels is given in run-length-encoding , one-encoding etc. More details are explained in the following sections.

4.1 Data Description and Data Pre-processing

4.1.1 Data description

The dataset under consideration in this study is a cloud satellite images, which was uploaded by [11] into Kaggle competition, under the name "understanding clouds in satellite images, the dataset was collected by NASA worldwide view. The data is attained from 3 various regions , of 21° longitude and 14° latitude spanning, the coloured images, whose are the original, were collected from Terra and Aqua, whose are two-orbiting satellites. Some images are combined from two different orbits and that's because of the small footprint of the mentioned satellites.

The uncovered area by the satellites is marked in black as it is shown in Figure 2.1. The dataset images consist of four different cloud patterns called Sugar, Flower, Gravel and Fish, these names were given to the clouds patterns by the scientist that labelled the dataset and for more information about the names and corresponding clouds patterns see section 2.2. The dataset consist of two files of images, training, and testing. The training dataset consist of 5546 × 4 and the testing dataset has 3698 images. The dataset images have the same size of 2100 x 1400 which was resized for pre-processing purposes. As it can be observed from Figure 4.2 ,4198 images has more than one label ,which makes it a multi-label and multi-class classification problem as well as , the sugar has the highest occurrences ,with occurrence of , in 2781 images and flower has least occurrence ,with occurrence of 2939 within the train dataset images. The training set was labelled by 68 different scientist and each image was labelled by three different individuals, more information about how the training dataset was labelled see [5].

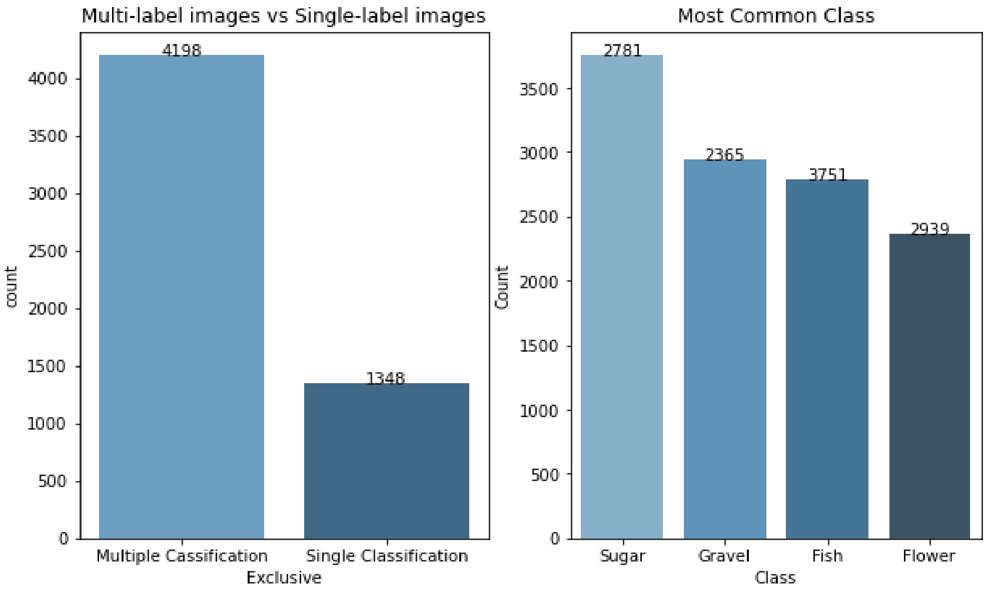


Figure 4.2: Images with multi-labels vs single-labels and Number of images per class

4.1.2 Data Pre-processing

This section will give brief overview of the data pre-processing techniques were used to prepare the ,in terms of the labels , the mask were given run-length-encoding values, and this was solved by decoding it , using python. Furthermore , the train datasets were augmented to increase the labelled data, and this was accomplished by using albumentations library from python ,which is explained in more details in 4.1.2.1 .The labelled images was split into train and validation 80:20 respectively.

4.1.2.1 Data Augmentation

Deep learning models consist of millions of parameters, for instance the models used in this study consist of more than 40 million parameters, these networks require a large, labelled data ,where in many cases is difficult to have such amount of labelled data to train deep learning model with large number of parameters. Therefore, the best solution to such problem is expanding the existing training labelled artificially by using a technique known as data augmentation. The technique is also known as class preserving transformation and it is applied in many state-of-the-art image's classification and segmentation models(Muaz, 2020) [36] . Hence, to increase the training dataset, which improves the models. Albumentations library was used, which efficiently transformed images using five different transformation. The images have been augmented into five types horizontal flip Figure 11.2,vertical flip Figure 11.3 , random rotation 90° Figure 11.4 , grid distortion Figure 11.5 and optical distortion Figure 11.6 .

4.1.2.2 One-hot encoding

Generally, datasets consist of categorical variables ,whether it's the dataset itself or labels of the datasets as for this dataset the class names are given in categorical variables, which makes challenging for the model to process. Therefore, it's very important the class names to be converted into numerical values. There are various techniques to convert the categorical variables into corresponding numerical such as label-encoding, custom-binary-encoding, one-hot-encoding and many others. However, in this study one-hot-encoding is used to convert the class's names from categorical to numerical as its one of the most used and efficient techniques.

4.2 Technical Approach

4.2.1 EfficientNet

One of the most advanced and recent modules introduced by google artificial intelligence researchers is EfficientNet ,which is a group of convolutional neural network with more improvement and higher performance than it is predecessors. The proposed network consist of eight models namely B0-B8 where each model differs from another with number of parameters and performance.

The EfficientNet operates in three techniques. The first way is depth wise + pointwise convolution, in this method performs in independent manner in each channel of the input, whereas the pointwise convolution combines the channel's output onto a new channels space, and it is 1x1 convolution. The second way is the "Inverse Res" which consist of two residual network blocks. The first squeezes the channels and the second extends it and links the skip connection to reach to the channel's layers. The third way that EfficientNet works is the bottleneck where it applies linear activation function, so that information will not be lost. The common models used to create all EfficientNet models are shown in Figure 4.3 , where module-1 is used as initial start point for all the sub blocks , module-2 is the starting point for all the sub-blocks excluding the first one , module three is used as connect skip connection across the sub blocks , module-4 is used for combining all the skip connection that happened in the first sub block and module-5 combines all the modules that are connected in skip connection in the previous sub-blocks. The modules are combined in different order to form the sub-blocks shown Figure 4.4 . EfficientNet is mainly built using Mobile-Net, which was originally used in MobileNetV2(Sandler et al., 2018) [39] , the basic architecture of EfficientNetB7 with its building blocks MBConv is presented in Figure 4.5 . All the 8 models of EfficientNet have common block with more complexity of architecture, for instance EfficentNetB0 consist of 237 layers in total, whereas EfficientNetB7 has 813 layers.

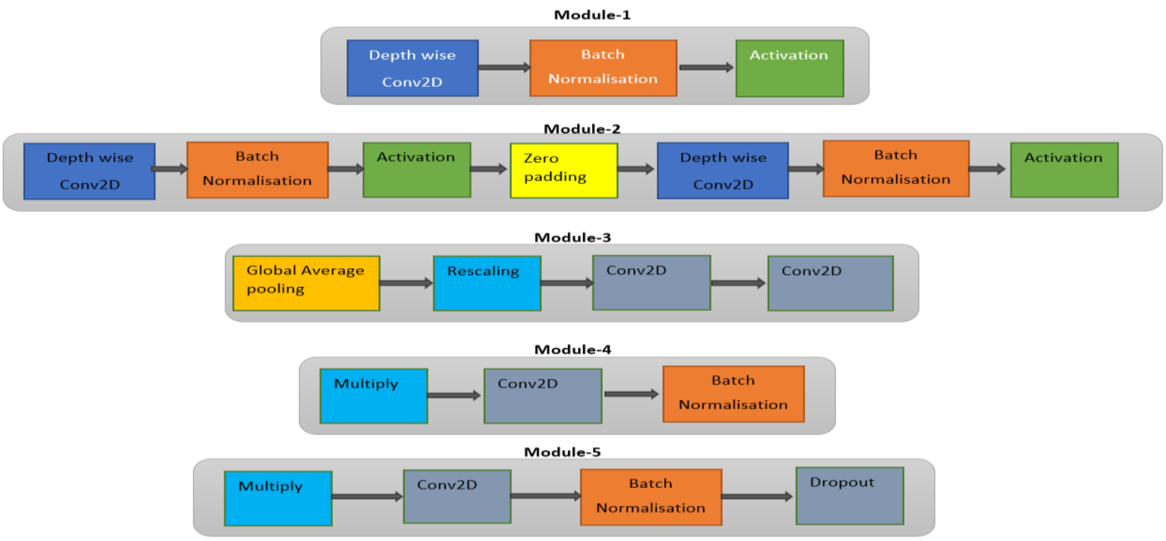


Figure 4.3: Common modules to form the architecture of EfficientNet

The sub-blocks formed by combining the modules in Figure 4.3 are used as follows. The first Sub-block is used as sub-block for the first block, the second sub-block is used as first sub-block in all the blocks and the sub-block is applied as a sub-block for all the block except for the block.

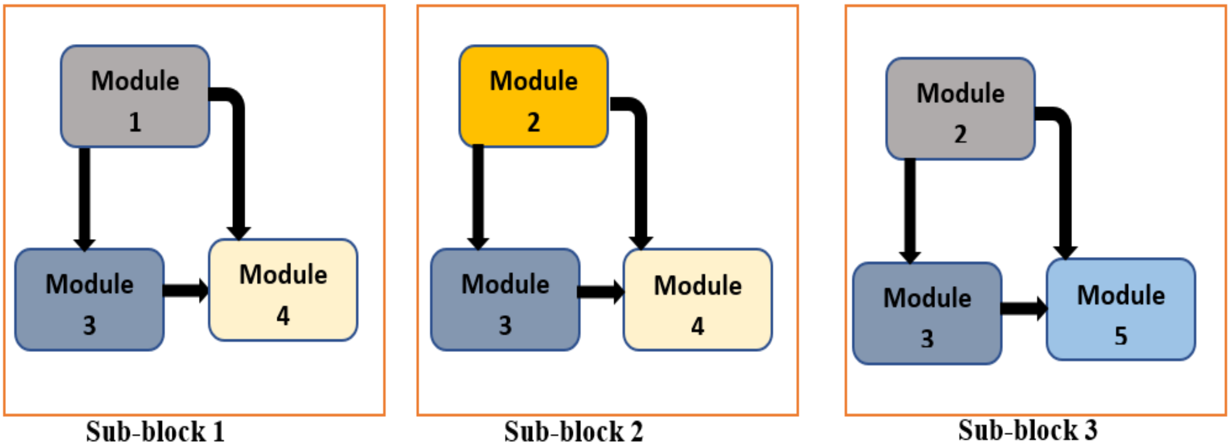


Figure 4.4: Sub-blocks with using individual blocks

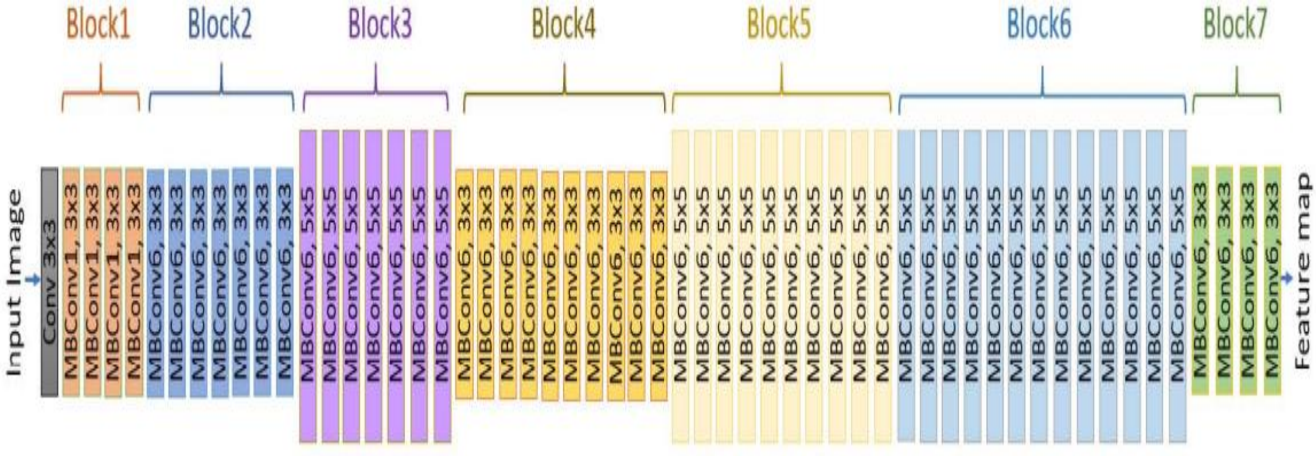


Figure 4.5: EfficentNetB7 architecture with MBConv as building blocks

The EfficientNet uses a scaling up technique known as compound scaling which was proposed in (Lee et al., 2020) [38]. The scaling technique is used to scale the dimensions, which is defined as systematic, and the principle factors of this scale are width scale, which adds more feature maps to all the network layers. Depth scale to add more layers to the model network and resolution. More details for the EfficientNet can be found in (Tan & Le, 2020) [37] . In this study, the EfficientNetB6 and EfficientNetB7 are prioritized to be used as a backbone or encoded for the models described 4.2.1.

4.2.2 U-Net

A U-shaped network was proposed by (Ronneberger et al., 2015) [40] under the name of Unet convolutional neural network for biomedical image segmentation. The network was modified and extended to operate with a smaller number of images and produce image segmentation with high precision. The main idea behind this network was to improve the encoder part by successive layers and further modification to network was replacing the pooling layer by upsampling operators. Therefore, the upsampling result in increased resolution of the network output. The upsampling part of the network have many channels, which is the main factor of the U-Net in propagating information to larger resolution layer. As the name defines the model consists of two parts , an encoder and decoder ,which form the U-shape of the network. The encoder part is normal CNN that have repeated convolutions, where each convolution is followed by an activation function (ReLU) and max pooling. During this process the spatial information is minimised, and feature maps are increased. The second part uses up-convolutions and concatenations to combine the feature and spatial information and give an output with high resolution. In another word the output size is same as the original input size.

4.2.3 Attention-Unet

The typical U-net discussed in section 4.2.2 , have two paths and it combines spatial information from the down-sampling part of the encoder with the up-sampling part to retain a good spatial information, in the process of reconstruction the original input size. However, this process passes poor feature representation from the initial layers of the encoder to decoder path. To overcome the limitations encountered by the original U-net and avoid bringing along poor feature representation to the decoder path, a new extension to the U-Net was proposed by (Oktay et al., 2018) [41].The architecture of the AttentionUnet is shown in Figure 4.6. The Attention U-Net input image is progressively filtered and at the same time, the input is down-sampled by a factor of 2 at the down-sampling or encoder stage. For instance, for Figure 4.6 , $H_4 = \frac{H_1}{8}$, where N_c denotes , the number of class. The network uses the Attention Gates (AGs) to filter the features and select the relevant features.

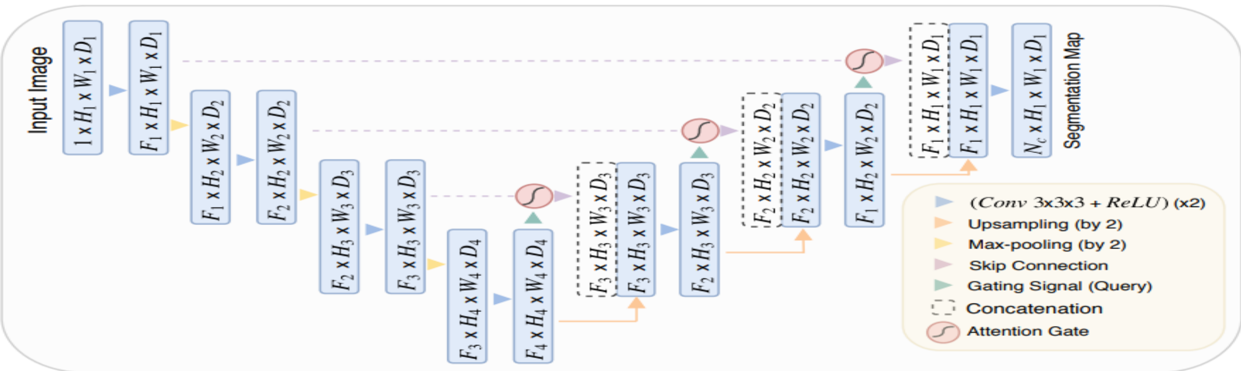


Figure 4.6: Attention-UNet

The AGs highlights the relevant features that are passed through the skip connection and within the skip connection the extracted information is applied in gating to disambiguate the irrelevant information or response in skip connection, and this is done before the concatenation is performed to merge only the relevant activation. Furthermore, the AGs filters the neurons as well as the irrelevant information from the background are downweighed. This allows the layers only to be updated with most relevant regions based on the task. The mathematical representation on how the convolutional parameters are updated in the layers are shown in equation (4.1).

$$\frac{\partial(\hat{x}_i^l)}{\partial(\Phi^{l-1})} = \frac{\partial\left(\alpha_i^l f\left(x_i^{l-1};\partial(\Phi^{l-1})\right)\right)}{\partial(\Phi^{l-1})} = \alpha_i^l \frac{\partial\left(f\left(x_i^{l-1};\partial(\Phi^{l-1})\right)\right)}{\partial(\Phi^{l-1})} + \frac{\partial(\alpha_i^l)}{\partial(\Phi^{l-1})} x_i^l \tag{4.1}$$

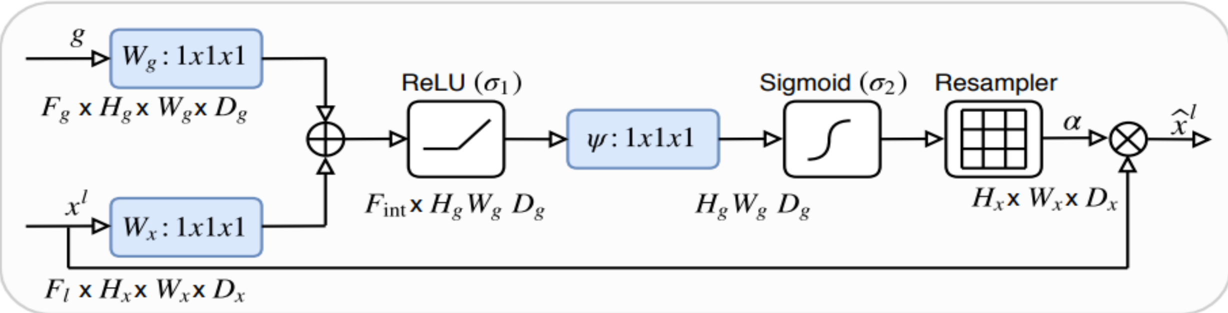


Figure 4.7: Schematic of additive attention gate

For better clarification of how the AGs works, lets break it down using AGs implemented in each skip connection shown in Figure 4.8. The AG takes two inputs, namely vector “X” and “g” , where “X” comes from the skip connection and it has better spatial information and “g” comes from the next lowest layer of the network and since it comes from deeper part of the network it has better feature representation. From Figure 4.8 the vector “x” has a dimension of (64x64x64) and the vector (32x32x32) which corresponds (filters, height, and width). The vector “x” is fed to a stride convolution and therefore its dimensions become (64x32x32) and the vector “g” goes through 1x1 convolution (64x32x32) and the two vectors are summed in element-wise which results in larger aligned weights and smaller unaligned weights .The output vector of the previous step is fed to a sigmoid layer and this layer scales the vector between the range [0,1] which produces the attention coefficients and the coefficient close to the value of 1 are considered more relevant features. Using trilinear interpolation, the attention coefficients are upscaled to the original dimension of vector “X” which is (64x64) and multiplied to the original vector “X” elementwise ,then fed to the skip connection.

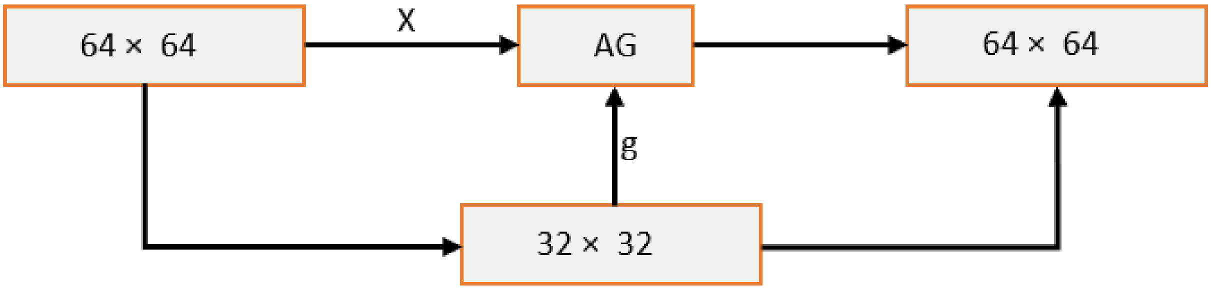


Figure 4.8: AG implementation breakdown for each skip connection

4.2.4 Residual Attention U-Net

Another deep learning algorithm for segmentation to overcome the limitations encountered by the original U-Net was Residual Attention U-Net(Chen et al., 2020) [42] shown in Figure 4.10 . Based on the problem on hand in this study a deep network is required, however the deeper the network the higher the layers is needed, which in so many cases this high number of layers will expose the network to converging problems such as when the network depth is increased the accuracy initially will be high and then decrease rapidly and this problem is defined as network degradation (He & Sun, 2015) [43] .To reduce the effect of degradation in the network while training (He et al., 2016) [44] proposed a technique that utilizes the skip connection with residuals and the technique helps to avoid estimating large number of parameters and avoid degradation. The proposed algorithm is known as ResNet and the typical block of ResNet is shown in Figure 4.9 A .

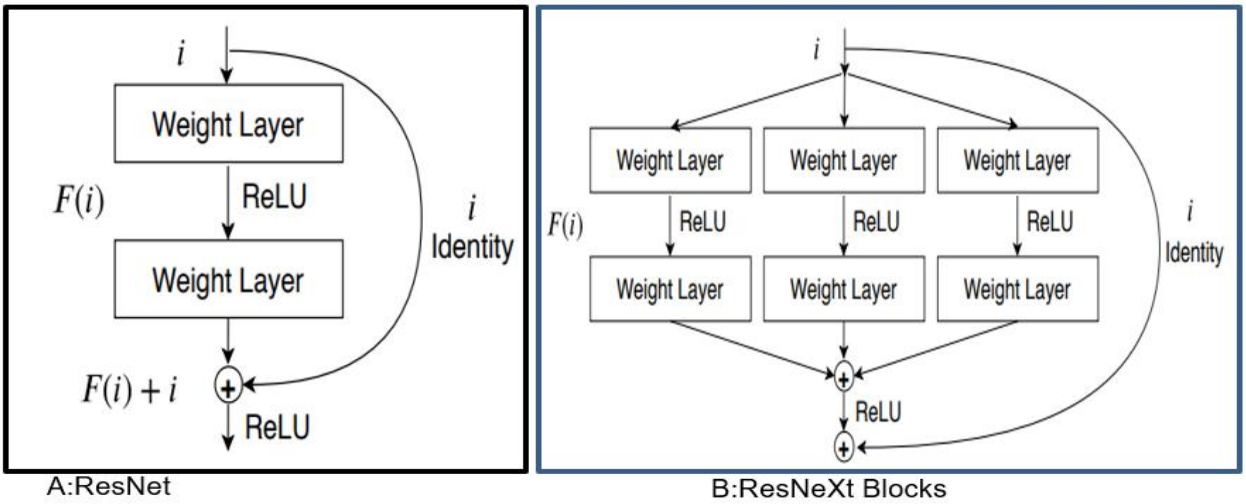


Figure 4.9:ResNet and ResNeXt block

From the ResNet block shown in the figure above, the i corresponds to D-dimension of the input image and with ResNet the skip connection is performed as identity mapping. The of the identity mapping output is summed with output of the stacked layers as follows.

$$F(i) = \sum_{j=1}^{D} w_j i_j \tag{4.2}$$

Where $i = [i_1 \dots i_D]$ and W is the trainable weight $W = [w_1 \dots iw_D]$ of the trainable layer. The ResNet differs from the U-Net by the features mapping in the decoding process, whereas the ResNet adds identity into each output block, and this allows the stacked residual block to have better learning of the latent representation of the input image. However , when increasing the number of layers, the ResNet becomes more complex and harder to converge and as solution to this problem Aggregated Residual Network (ResNeXt) was proposed by (Xie et al., 2017) [45], where this network increases the cardinality instead of depth ,which is shown in Figure 4.9 B and it is mathematical formulation is :

$$F(i) = \sum_{j=1}^{C} T_j(i) \tag{4.3}$$

C corresponds to number of residual transformations, which will be aggregated and by considering a simple neuron T_j is transformation resulting i in low-dimensional embedding and then will be transformed and the extended residual function can be represented by equation (4.4)

$$y = \sum_{j=1}^{C} T_j(i) + i \tag{4.4}$$

Where y corresponds to the output, in general the ResNeXt has small difference from the typical ResNet by the size of the weight of the layers and this because the ResNeXt reduces the layers by using cardinality and keeps the same performance. One important thing about the ResNeXt,is the residual blocks should have the same topology. The ResNeXt blocks are featured within Residual Attention U-Net shown in Figure 4.10 to capture features from the original input image , the developed network works as follows , the decoder path receives features learnt the encoder path and concatenate the output of deconvolutional layers , which is followed by Attention Gates mechanism similarly to the one discussed in section 4.2.3. The network is used for semantic segmentation and its performance will be compared with other networks used in this study.

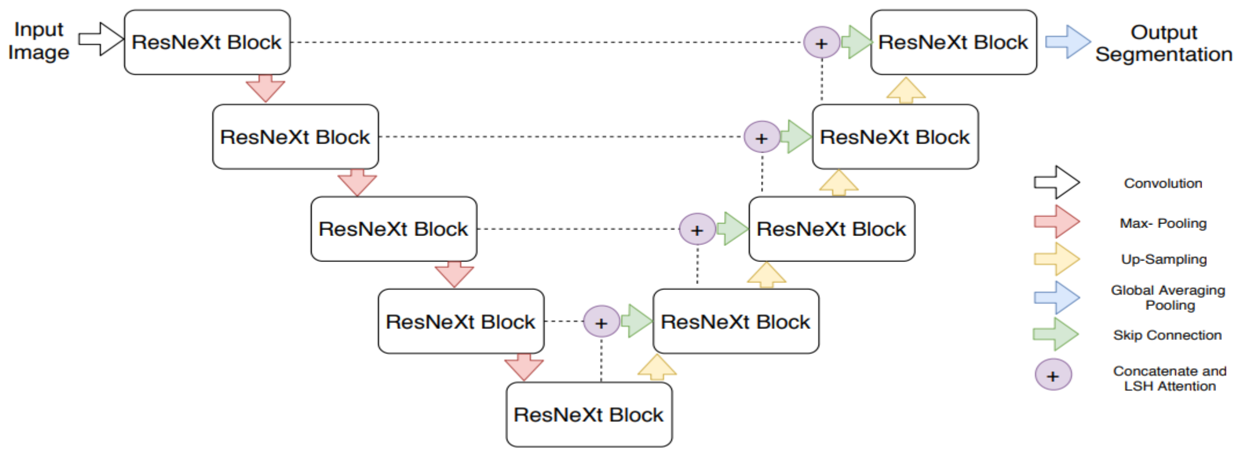


Figure 4.10: Residual Attention U-Net architecture proposed by[42]

4.2.1 EfficientNet-Unet

The encoder part in the conventional, is close to symmetrical to the decoder. Whereas in this study, the EfficientNet is used instead as encoder in the contracting path. The decoder or expansion is like the decoder from the Unet model. The model architecture is illustrated in Figure 4.11 considering EfficientNetB7 as encoder which is shown in Figure 4.5. The original image size of the dataset under consideration in this study is 2100x1400, however, it was resized to 128x256 due to computing power limitations and further processing. The working process of the model is as follow, feature maps are unsampled bilinearly of the last logit of the encoder, the upsampling factor is two and then concatenates it , with the first part feature map, and the resolution spatial is the same ,then it is followed by 3x3 convolutional layers and then again upsmapling is performed by factor of 2. This operation is repeated until segmentation map of same size as the original input image is reconstructed.The model has asymmetric architecture unlike original U-Net. This study prioritised EfficientNetB6 and EfficientNetB7 as encoders to form the EfficientNet-UNet.

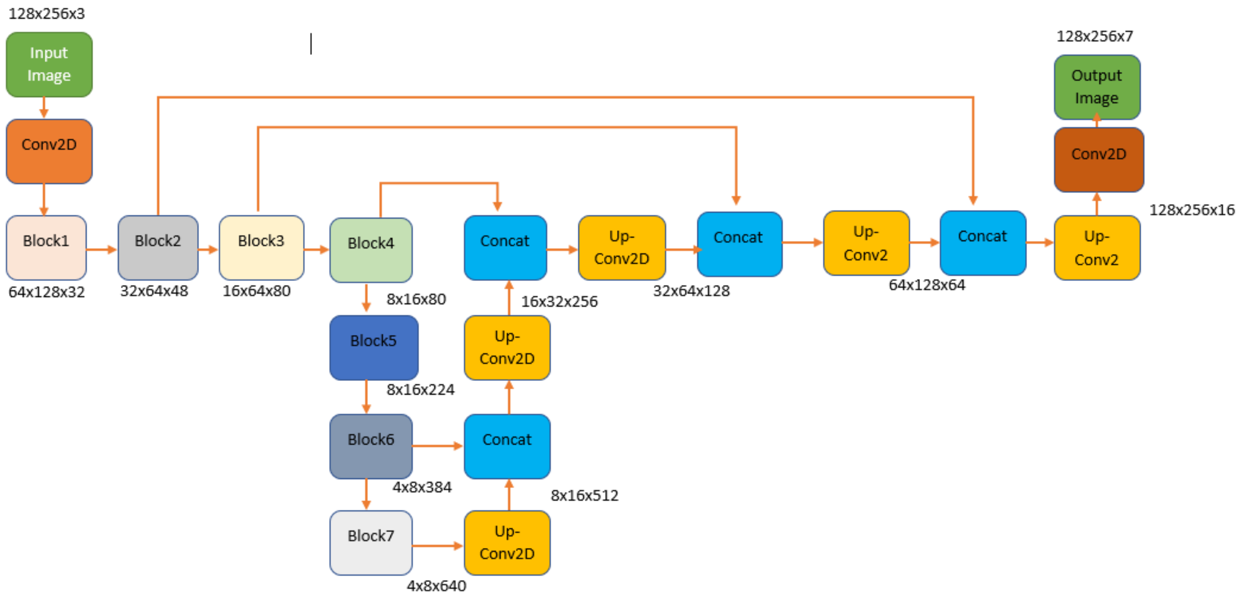


Figure 4.11: EfficientNetB7-Unet

5 Design

This section will present the design process of the models discussed in 4.2 , the deep learning implementation for this task can broke into four distinct steps.

- Data pre-processing: loading dataset and conducting data preparation discussed in section 4.1.2.
- Model training: Create the segmentation models presented in section 4.2 and train the models on the labelled dataset.
- Model evaluation: Evaluate the models, using the evaluation metrics presented in section 7.1.
- Prediction: Make masks prediction using the trained models on unseen cloud images data.

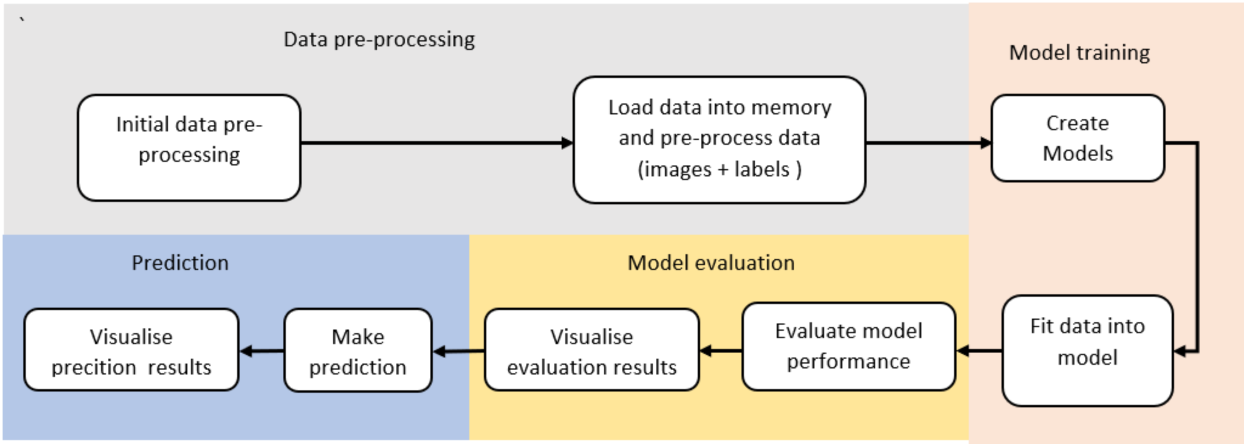


Figure 5.1: Flow chart of clouds classification using semantic segmentation models to implement

5.1 Technical design approach

5.1.1 Optimiser

An optimiser is an algorithm that is used to change the attribute of a model such as weights and learning rate to minimise the losses and it is uses minimising functions to solve an optimisation problem. Optimiser is also responsible for providing the most accurate results and that is achieved by reducing the losses. Various optimisers were proposed and studied by researchers over the years such as Gradient Descent, Stochastic Gradient Descent, Mini-Batch Gradient Descent, Momentum, Nesterov Accelerated Gradient, Adagrad, AdaDelta, Adam and Adam, each optimiser has its advantage or disadvantages. However, RAdam produced the best results, therefore, it was used as an optimiser for all the models. The RAdam optimiser was proposed (Liu et al., 2020) [49], and it was introduced to overcome the problem encountered by adaptive learning, which is the large variance on the early stage of the model training and this because it uses limited amount the training samples, and to solve this problem and this was solved by using warmup heuristic and the RAdam operates as variance reduction technique .It is suggested to start with very low learning rate as a warmup.

5.1.2 Loss optimisation

As it is known deep learning algorithms applies optimisation functions to optimise and learn objectives. However, for the purpose of learning an objective in a very fast and accurate manner, we need to ensure the coverage of even the edge cases by our objective's mathematical representation or loss. The history of loss functions have roots in traditional machine learning, the loss functions were derived based on label distribution, for instance , binary cross entropy was derived from Bernoulli distribution, whereas categorical cross entropy was derived from multinoulli distribution (Jadon, 2020) [50]. There are various loss optimisation function, however, for this two-loss function were used Jaccard loss and dice loss.

5.1.2.1 Jaccard Loss

The jaccard index or intersection over union discussed in section 7.1.2 , it also has corresponding loss, which was proposed by (Berman et al., 2018) [53], the proposed jaccard loss can be formulated mathematically as follows:

$$\Delta_{J_c}(X.Y) = 1 - \frac{|X \cap Y|}{|X \cup Y|} \quad (5.1)$$

5.1.2.2 Dice Loss

One of the most known metrics in the deep learning community is dice coefficient , where it measures the similarity between two images , this technique was adapted as loss function by (Sudre et al., 2017) [51] and it is called dice loss. Dice loss tries to increase the overlap between the ground truth pixel value and the prediction.

$$Dice\ loss\ (y, \hat{p}) = 1 - \frac{2y\hat{p}+1}{y+\hat{p}+1} \quad (5.2)$$

The reason 1 is added in both denominator and numerator is to ensure the function is not undefined in edge cases scenarios such as $\hat{p} = y = 0$.In this study , dice loss was combined with binary cross entropy.

5.1.3 Learning Rate

The learning rate can be defined as hyperparameter that controls the model change in response to the error to each time the model weights are updated. If the leaning rate is too small model takes too long to converge and it is ideal to start with acceptable initial learning and then updating it using a decay factor, in this study plateau is used to change the learning based on the model weights update. The learning rate can be represented mathematically as follow.

$$\epsilon_t = \epsilon_0 \forall_t < \tau \quad (5.3)$$

$$\epsilon_t = \epsilon_0 t^{\alpha} \quad (5.4)$$

5.1.4 Regularisation

During the training process an early stop is used, and the reason early stopping is used is, to ensure the model generalises good for the testing data. The validation set is used to make predictions at the end of each training epoch and that is achieved by computing the loss and accuracy. Furthermore, the validation loss is monitored for the purpose of early stop and if there is not improvement in the validation loss, the model will stop training before reaching the maximum epoch and that will prevent model overfitting. Another technique used to generalise the models for the unseen data and avoid overfitting is dropout, which is discussed in section 3.2.5.1. A dropout of 0.2 was used with all the models excluding the EfficientNet-Unet models.

5.2 General design decisions

5.2.1 Programming language

Many programming languages can be used to implement this project such as R, MATLAB, Java, Java script. However, python was chosen due to familiarity and experience with the programming language and many other factors such as the availability of many options of free open-source libraries for implementing deep learning functionality. The main libraries used in this project are TensorFlow, Keras, Matplotlib, albumentations, Numpy, tqdm ,CV2 , Pandas and SciKit-Learn.

5.2.2 Deep Learning Framework

Due to the nature of the dataset, which is images dataset and complexity of the deep learning to implement, a powerful computing is required, therefore, for this project, a laptop with Graphical Processing Unit (GPU) NVIDIA GEFORCE GTX 8GB and core i7 processor is used. Cuda is also used to make use of the GPU’s computing capabilities, because deep learning works with CUDA for parallel computing and GPU optimisation.

5.2.3 Interface

Jupyter notebook was chosen as interface for this project due to its functionality such as interactive computing, kernel as computing, clear results and visualisation and markdown such as headings and commenting.

5.2.4 Design decisions Summary

This section summarise the design decision for this project as follows.

1. Dataset:
    - Understanding clouds from satellite images which a multi-class and multi-label classification.
2. Data pre-processing:
    - Split labelled dataset into 80:20 training and validation.
    - One-hot-encoding for class names.
    - Decode the masks from run-length-encoding.
    - Use albumentations library for data augmentation for the purpose for increasing labelled dataset.
    - Image normalisation.
3. Model training :
    - Segmentation models: U-Net , Attention_Unet, Residual Attention U-Net and EfficeintNet-Unet.
    - SoftMax as output layer activation function.
    - Optimiser: RAdam
    - Loss functions: Jaccard and dice loss.
    - Generalisation and regularisation: Dropout, early stopping.
4. Evaluation metrics:
    - Dice Coefficient.
    - Jaccard Index or Intersection over union
5. Programming language and open-source frameworks:
    - Python 3.8
    - TensorFlow 2.5 and keras, Numpy, Matplotlib, Pandas, CV2, Albumentations and Scikit-Learn.
6. Interface: Jupyter Notebook

6 Implementation

6.1 TensorFlow

All experiment on this study were carried out on TensorFlow 2.5.0 . TensorFlow is open-source s library developed by Google artificial intelligence team, the first version of TensorFlow was released in 2015, which is TensorFlow 1.0, and it allows users to perform arbitrary computation as s graph of data flows. The nodes within this graph is the representation of the mathematical operations and the edges are the representation of the data communicated between one node and another. Tensors, which are multi-dimensional arrays, are the representation of data in TensorFlow. TensorFlow is mainly used for deep learning research and practice (Zaccone et al., 2017)[47] . Even though TensorFlow is powerful, it was not always the first choice for researcher. However, the TensorFlow team started addressing this issue and released TensorFlow 2.0 which is more stable and intuitive. The major changes was integrating keras, which is open-source neural network library that is user-friendly written in python (Lee, 2019) [48] .

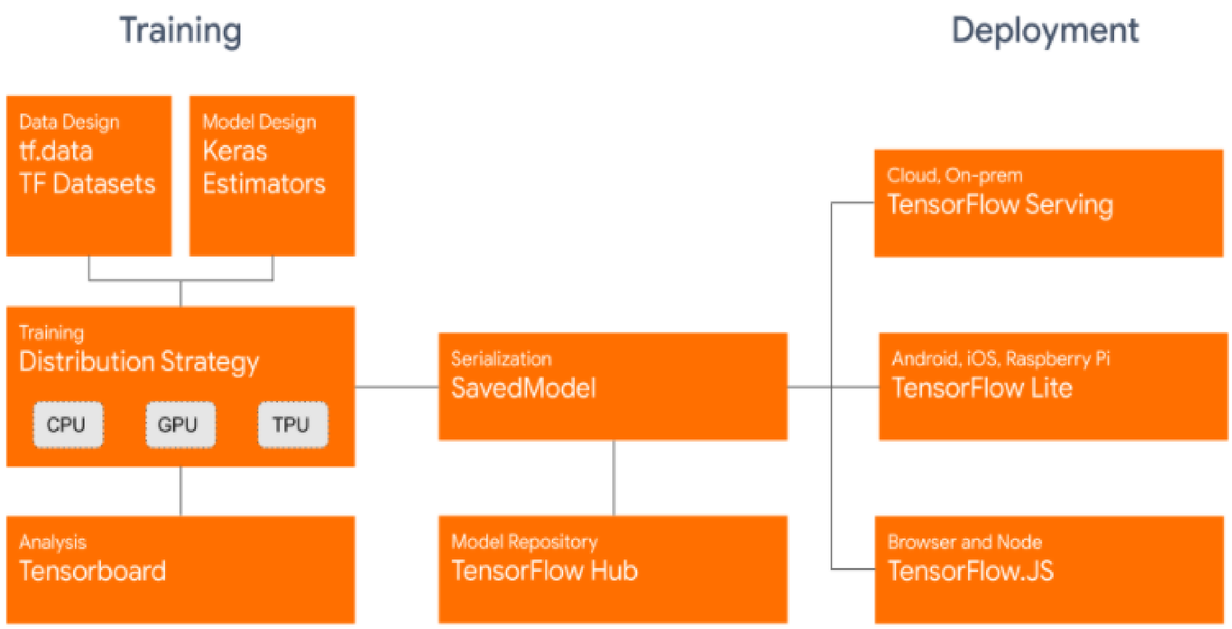


Figure 6.1: TensorFlow

6.2 Code Design

For reproducibility purposes this section will give a brief explanation of the experiment's implementation in terms of coding as well as the flow of work. The implementation was conducted in three section in terms of coding the first part is loading the dataset and conducting exploratory data analysis and initial data visualisation. After initial data pre-processing. Furthermore, pre-processing and data preparation, this includes data augmentation, normalisation and generating training, validation and testing instances was executed inside a class named Data Generator. The purpose of using the Date Generator is to load the date in sequence to avoid running out of memory while loading the dataset. The keras data generator class. The Data Generator adapted for this study consist of two methods whose are fit_generator and predict-generator. The fit-generator has two generator one for training and another one for validation, both generators return inputs and targets and both are instance of sequence class. The predict-generator returns only the inputs in another word the “X”.

6.2.1 Models and Weights Saving

Once the model is trained, the weights and the model are saved to be re-used for further evaluation transfer learning experiments, this is achieved by saving the model in HDF5 format. The saved model includes the model configuration, weights, the optimisers, the training results. Within the training it is an automated task to save the best model and for further understanding this can be observed in the in link for source code provided in **Appendix F** .

6.2.2 Prediction and results visualisation

When the model training is finished, predictions can be performed on the unseen, and this can be achieved by using the data generator and generate input instance and the pass the input instance and their ground truth labels. Once the prediction is complete, the prediction results can be visualised as such original with original mask alongside the original image with predicted mask. Sample of the prediction results is presented in Appendix E.

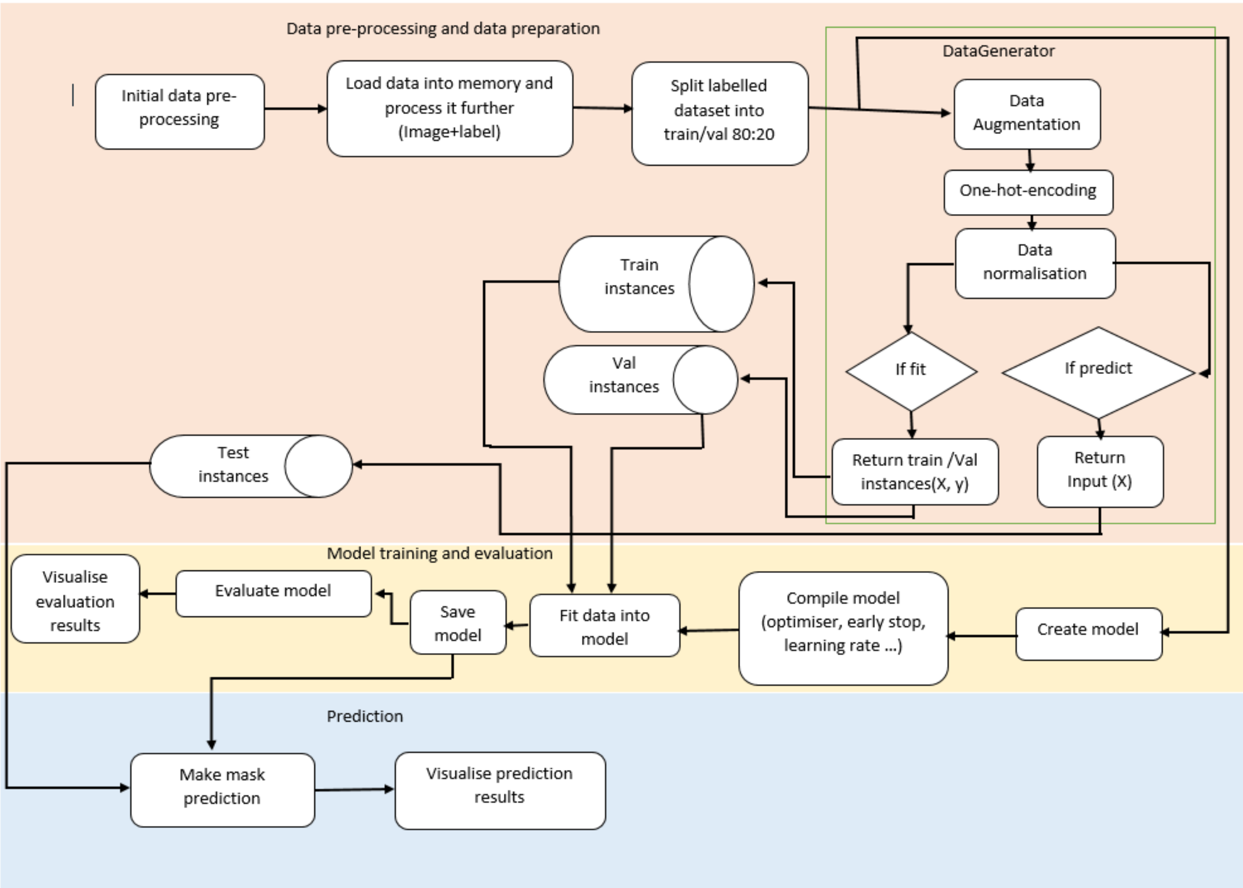


Figure 6.2: Detailed implementation flowchart

7 Results and Evaluation

7.1 Evaluation Metrics

As semantic segmentation been prioritised in this study , for the purpose of segmenting each region of the image belonging to specific class of the four clouds classes using deep learning models. Hence , it is very important to evaluate the model's performance on the underlying dataset and determine its suitability for the application requirement. Therefore, this section presents the evaluation metrics used to evaluate all the models.

7.1.1 Dice Coefficient

The dataset provider recommended the use of dice coefficient as an evaluation metrics. Therefore, this metric was used as an evaluation metrics to evaluate all the models. It was applied to compare the pixel similarity between the precited segmentation results and the ground truth. Dice coefficient can be formulated mathematically as follows (Dice, 1945) [52].

$$Dice\ Coefficient(DC(Y,X)) = \frac{2\times|X\cap Y|}{|X|+|Y|} \tag{7.1}$$

Where X corresponds to the predicted number of pixels .whereas Y represents the ground truth .Dice coefficient ranges between 0 and 1 , where 1 being the highest similarity between the predicted and ground truth and 0 the opposite. It is also different from other evaluation metrics such as accuracy because it is used to validate segmentation techniques performance and it is achieved by computing the dice score, which is the 2 multiplied by the of overlap between the two images divided by the total area of both images.

7.1.2 Jaccard Index or Intersection over Union (IoU)

Another evaluation metrics that was used in this study was the Jacard Index also known as intersection over union (IoU),this metrics , is one of the commonly used evaluation metrics in semantic segmentation and it is straightforward, and it is simply the area of overlap between predicted segmentation results and the ground truth divided by the union of predicted area the ground truth area, similarly to dice coefficient it ranges between 0 and 1, with 1 indicating the highest overlap between prediction ground truth. The mathematical formulation of jacard index is as follows.

$$IoU\ or\ Jacard\ Index = \frac{|X\cap Y|}{|X\cup Y|} \tag{7.2}$$

Or using the dice coefficient mathematical formulation , the intersection over union can be represented mathematically as follows.

$$IoU\ or\ Jacard\ Index = \frac{DC(Y,X)}{2-DC(Y,X)} \tag{7.3}$$

7.2 Results

This section of the report presents the results of clouds classification in satellites images using segmentation models, all the models were trained in the same manner and same training parameters were used. The parameters used for the training are shown Table 7.1.The maximum training epoch is 30, however, early stop was used to stop training the model to avoid overlap if there is not improvement in the validation loss. Another technique used to improve the model performance while it is training is by reducing learning rate with patience of 3, in another word, the learning rate will be reduced after 3 epochs. This will allow to make changes to improve model performance while it is training. The total number of models parameters, trainable and non-trainable are shown in Table 7.2 , with millions of parameters for each model.

Table 7.1: Training Parameters

Warmup learning rate	0.0002
Early stop patience	5
ReduceLROnPlateau patience	3
Decay drop	0.5
Maximum Epoch	30

Table 7.2: Models number of parameters

Model	Total parameters	Trainable parameters	Non trainable
U-Net	31,402,708	31,390,924	11,784
Attention U-Net	37,334,872	37,319,248	15,624
Residual Attention U-Net	39,090,584	39,069,072	21,512
EfficientNetB6-UNet	37,435,258	37,227,151	208,107
EfficientNetB7-UNet	53,564,192	53,288,201	275,991

7.2.1 Unet Results

Figure 7.1 illustrates the training and validation of Unet with dice coefficient and loss as well Jaccard Index and loss. With dice coefficient and loss , we observe from the obtained results that the model has high divergence and keeps training for 17 epochs before its stabilised and stop training due to unnoticeable improvement in the validation loss, With the jacard index and loss the model is trained to the maximum number of training epochs ,which is 30.However , before stopping training the model starts to diverge ,which is unappreciated.

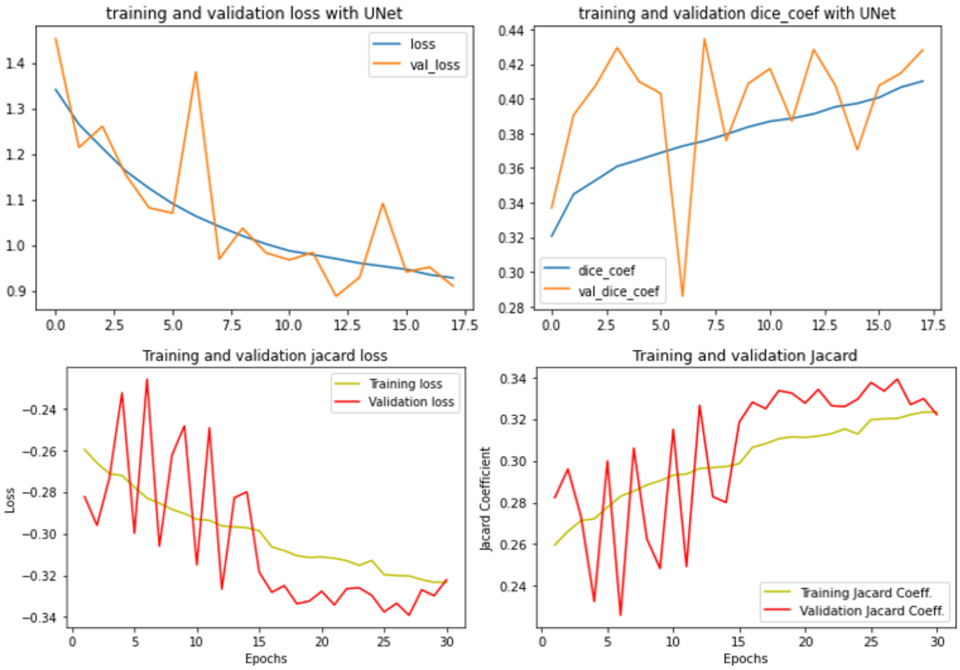


Figure 7.1:Unet dice loss, dice coefficient , Jaccard index and Jaccard loss for training and validation

7.2.2 Attention U-Net Results

The same process of training was followed to train Attention U-Net, and the results are shown Figure 7.2 ,and from the obtained results ,it can be observed that the model took more epochs to train before it stops training , which is 23 epochs. However , an improvement can be seen in the validation data. With Jacard Index ,the model continued to train to the maximum number of epochs and the results shows very stable improvement in the model training performance after one divergence occurrence ,which is a drop in the Jacard index score.

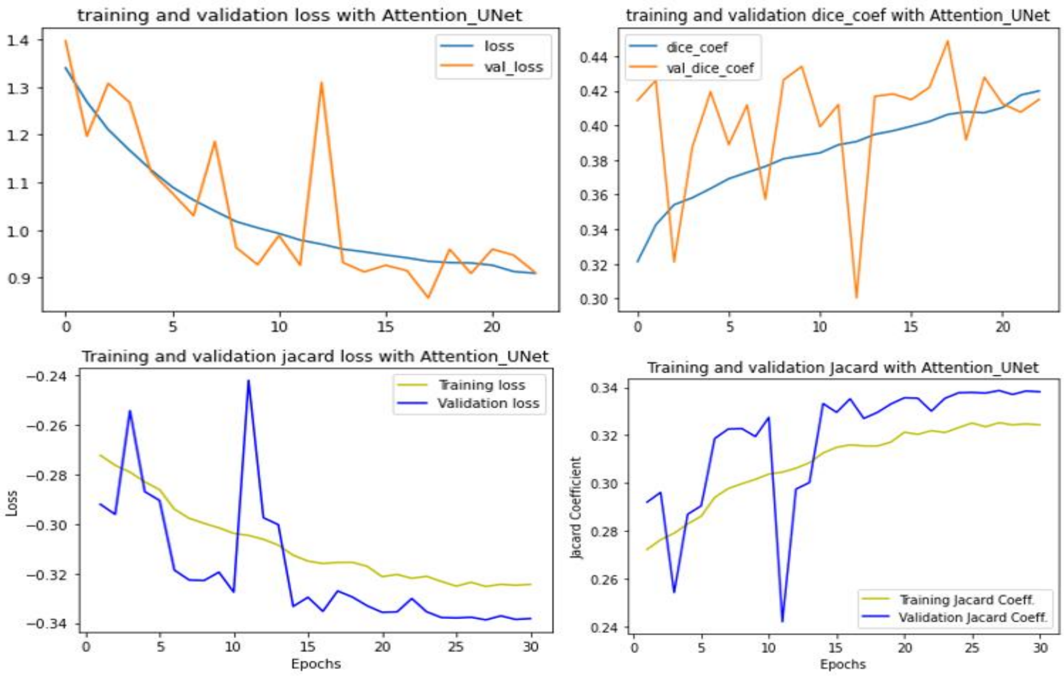


Figure 7.2 Attention U-Net dice loss, dice coefficient , Jaccard index and Jaccard loss for training and validation

7.2.3 Residual Attention U-Net Results

As it can be seen from the results illustrated in Figure 7.3 during training residual Attention U-Net , the model kept training to maximum training epoch in the first phase which is training with dice loss and the results shows , a slight improvement of model performance in comparison with the previously discussed is seen.

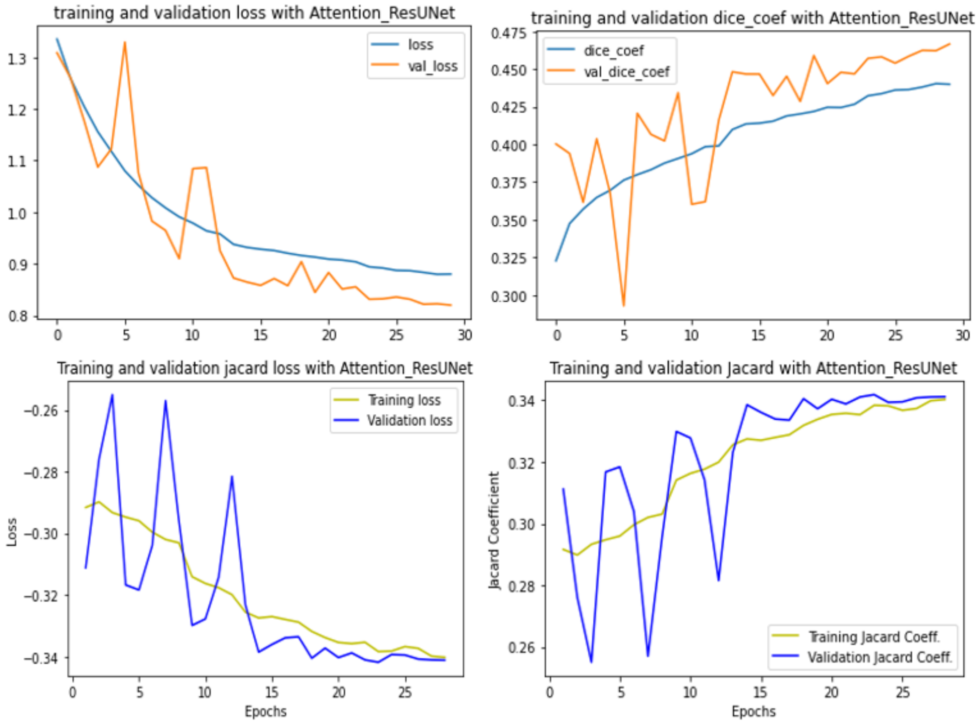


Figure 7.3:Residual Attention U-Net dice loss, dice coefficient , Jaccard index and Jaccard loss for training and validation

7.2.4 EfficientNetB6-UNet Results

Using the same procedure EfficientNetB6-UNet was trained and evaluated and from the attained and presented in Figure 7.4 , we observe that, the model has significant improvement in comparison with the other models and the dice score progressively improved throughout the training duration, With both losses , dice loss and jaccard loss , the model did not reach the maximum training epochs. Similarly , EfficientNetB7-Unet results is presented in Figure 7.5, we observe that , the model has, a heavy oscillation throughout the training duration ,even though the performance improved in the final epochs of the training and this could be due to amount of parameters, that the model has. In terms of loss, we observe that the model has big loss at the start of the training, but it drops rapidly to acceptable range ,which reflects positively in the model performance. The model's results are compared and discussed in more details in 7.3.

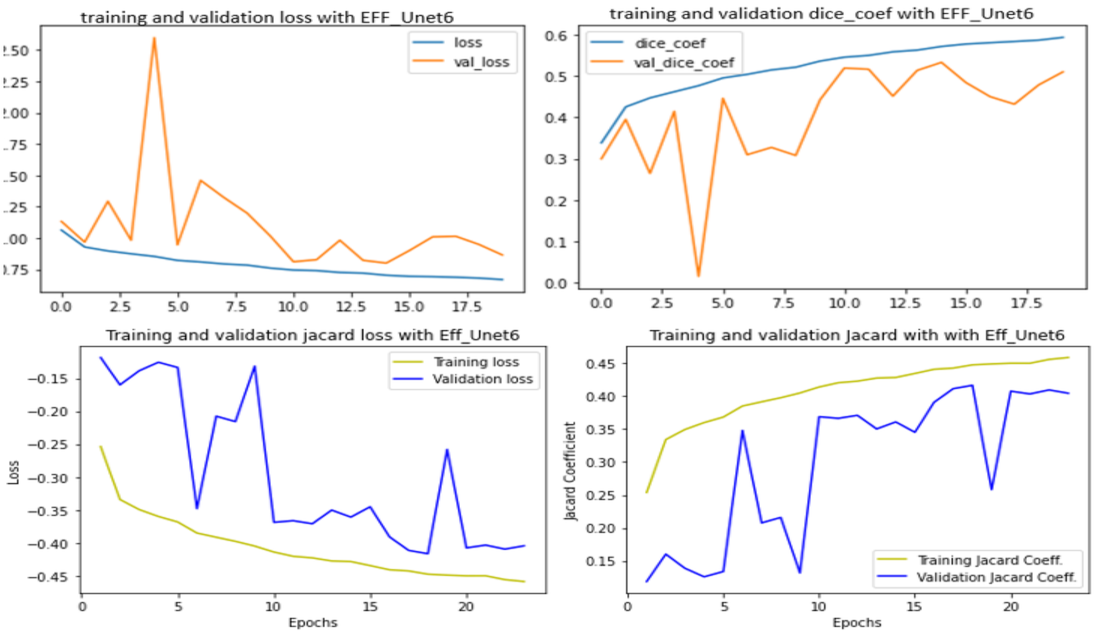


Figure 7.4: EfficientNetB6-UNet dice loss, dice coefficient , Jaccard index and Jaccard loss for training and validation

7.2.5 EfficientNetB7-UNet Results

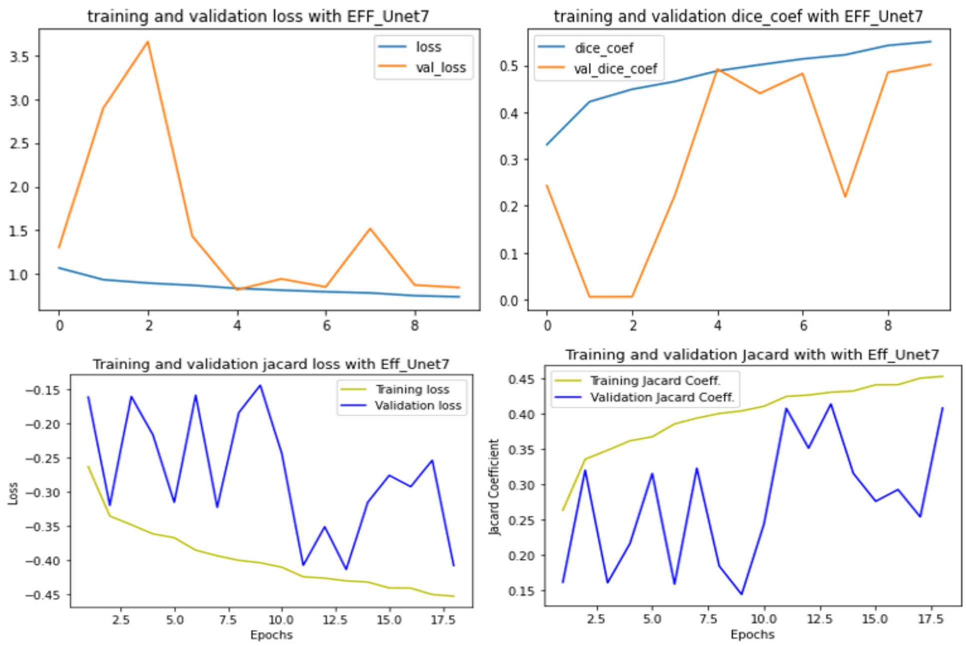


Figure 7.5: EfficientNetB7-UNet dice loss, dice coefficient , Jaccard index and Jaccard loss for training and validation

7.3 Models Comparison and Discussion

Table 7.3: Models dice score and IoU in training and validation set

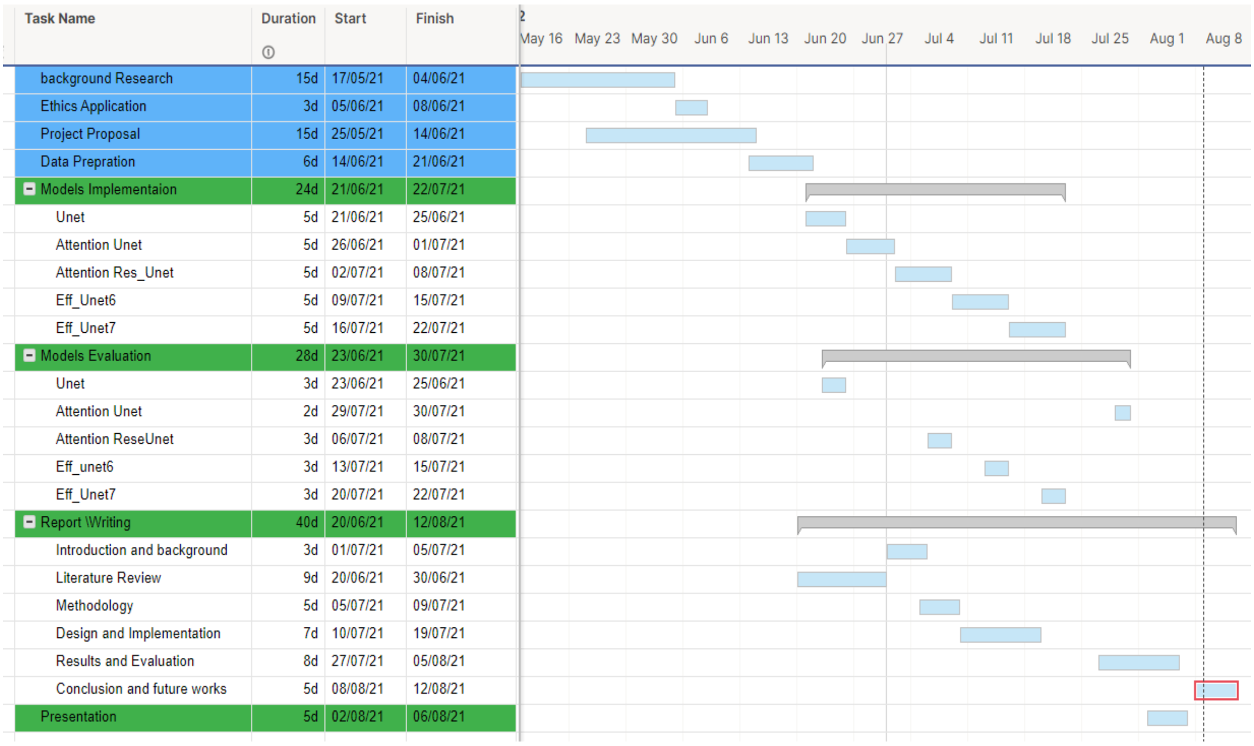
Model	Dice Coefficient		Intersection over Union	
	Training	Validation	Training	Validation
U-Net	0.4301	0.4282	0.3233	0.3221
Attention U-Net	0.4463	0.4488	0.3382	0.3243
Residual Attention U-Net	0.4600	0.4667	0.3403	0.3412
EfficientNetB6-UNet	0.5934	0.5329	0.4587	0.4045
EfficientNetB7-UNet	0.5503	0.5014	0.4531	0.4083

From the obtained results for all five segmentation models used in this study for classifying clouds structure into four different classes, it was possible to analyse and compare model's suitability for the application of clouds classification. As it was discussed in section 4.1.2 , the original images size is 2100x1400x3 and it was resized to 128x256x3 due to computing power limitation. Even though , the images were resized , the segmentation model showed a potential of success in this dataset , which indicates the limitations is not on the deep learning but on the computing power and data availability. From the evaluation metrics results obtained and presented in Table 7.3 ,it can be seen that the performance of the models varies ,with EfficientNetB6-Unet achieving the best performance among the models ,on both training and validation with dice score of 0.59 and 0.53 in training and validation respectively as well as IoU of 0.46 and 0.40. Followed by EfficientNetB7-UNet. As it was discussed in section 4.2 ,that Attention U-Net was proposed to overcome the limitation encountered by U-Net and similarly , Residual Attention U-Net to overcome the limitations of Attention U-Net ,the obtained results proved this theory ,as it can be seen , Attention U-Net has better performance than U-Net , however Residual Attention U-Net performed better than U-Net and Attention U-Net. The results shows the gradual trend of improvement on the models' performance and it indicates that , it is possible to achieve good classification results with used of good encoder such as EfficientNet. The encoder can be replaced by any model.

8 Project Management

8.1 Project Schedule

To be able to achieve this project goal , a project schedule was created in the very first week, which is the Gantt chart shown in the figure below , the project was divided into tasks ,whose are background research , Ethics application , project proposal, data preparation , models implementation, models evaluation , report writing and presentation. The project plan also shows the allocated time to each task. This plan highly contributed on completion of this project within the specified timeframe, and it was possible to meet all deadlines.



Task Name	Duration	Start	Finish
background Research	15d	17/05/21	04/06/21
Ethics Application	3d	05/06/21	08/06/21
Project Proposal	15d	25/05/21	14/06/21
Data Prepration	6d	14/06/21	21/06/21
Models Implementaion	24d	21/06/21	22/07/21
Unet	5d	21/06/21	25/06/21
Attention Unet	5d	26/06/21	01/07/21
Attention Res_Unet	5d	02/07/21	08/07/21
Eff_Unet6	5d	09/07/21	15/07/21
Eff_Unet7	5d	16/07/21	22/07/21
Models Evaluation	28d	23/06/21	30/07/21
Unet	3d	23/06/21	25/06/21
Attention Unet	2d	29/07/21	30/07/21
Attention ReseUnet	3d	06/07/21	08/07/21
Eff_unet6	3d	13/07/21	15/07/21
Eff_Unet7	3d	20/07/21	22/07/21
Report \Writing	40d	20/06/21	12/08/21
Introduction and background	3d	01/07/21	05/07/21
Literature Review	9d	20/06/21	30/06/21
Methodology	5d	05/07/21	09/07/21
Design and Implementation	7d	10/07/21	19/07/21
Results and Evaluation	8d	27/07/21	05/08/21
Conclusion and future works	5d	08/08/21	12/08/21
Presentation	5d	02/08/21	06/08/21

Figure 8.1: Project Gantt Chart

8.2 Risk Management

For the completion of this project, it is essential to identify the potential risks can be encountered and disrupt the progress of the project , therefore all the potential risks were identified, and possible solution was proposed. Hence , identifying them beforehand is very important for the purpose of completion of this project. From the risks presented in Table 8.1,I was exposed to the risk number 1 and 2. The first risk ,when all my university accounts were blocked ,that includes my one drive, where all my work is saved .However , I was able to continue the project because of the regular back of work in different locations. The second risk , difficulty of images processing due to large size of the images of this study dataset, and it was solved by resizing the images and using smaller batch size.

Table 8.1: Risks associated with project

Number	Risk	Risk Reduction	Solution	Impact of risk
1	Loss of dataset	Save the data in one-drive . hard drive and google drive	The dataset was downloaded from publicly available source, however ,in some cases the data can be removed or modified	High
2	Technical difficulty of processing image with large size	Resize images if possible and use small batch size	Cloud computing can be used such as google collab , however in this study it found slow or use high power computing if it can be provided by university.	Medium
3	Loss of tools or equipment	Save work regularly	The only that was used in this project was a laptop and it is good to have a backup computer	Medium
4	Health and safety risks	Work from home if possible	Due to the Covid-19 ,working from home and managing social distance to avoid being exposed to the virus.	High

8.3 Quality Management

To maintain the quality of this project within the specified timeframe ,various techniques was used. The main one being , consistent meetings with supervisor . to discuss results , doubts and set next steps. Another technique was to follow the plan that was created for the project. The meetings records with supervisor can be found in **Appendix A**. This action helped to maintain the quality of this project.

8.4 Social, Legal, Ethical and Professional Considerations

The dataset is publicly available on Kaggle competition at no cost ,however , and it is free from sensitive factors that could affect anyone, and the legal ownership of the data used in this project was acknowledged and project was approved, and the ethical certificate of the project can be found in **Appendix C**.

9 Critical Appraisal

Understanding cloud's structure is not a new field of study , As it was discussed in the literature review , classifying clouds patterns is very important to many applications and with accurate detection and classification of clouds beforehand ,could help to save lives and millions of dollars. One of the main contributions of this project is the usage of segmentation models in classification task ,for the purpose of classifying each region of the image belonging to specific class of the four cloud classes and that was achieved by using deep learning models.

One of the problems that was analysed in this project is the complexity of the dataset and similarity between the four classes in terms of patterns, however , the models using good encoder such as EfficientNet showed the potential of successful clouds classification. With all the challenges encountered the models shown a promising result and it can be improved.

10 Conclusion

This project proposed and implemented classification of cloud structures using satellite images and deep learning segmentation models into four different classes ,whose are sugar , flower , gravel, and fish. Five different deep learning segmentation models were used, and the performance was compared using two different evaluation metrics, dice coefficient and jaccard index which is also known as intersection over union (IoU).The project has proven with the use if deep learning segmentation models, specially encoder-decoder network and adapting a good encoder such as efficientnet, this will result in good performance in this dataset and perfectly classify cloud structures, which helps in interpreting clouds formation pattern, despite the obtained results which is less appreciated due to computing power limitation. As it has been proven ,the encoder path of the U-Net can be replaced by any other scaled up neural network and boost the performance , however , a model complexity was encountered , which means a bigger model has more parameters to train and this requires more dataset and computing power.

10.1 Achievements

The main objective of this project was to classify cloud structures into four different classes and as deep learning was concluded from the literature review ,that it has the potential to achieve the objective. However , throughout the project there were limitations encountered that affected the project outcome , the main limitation is computing power , due to the bigger size of the images and complex architecture of the models ,that requires higher computing power to train the models with the original images size. Despite all this limitation, the project achieved the main goal and implemented a new technique of classifying cloud structures and that is by using segmentation models. Even though the results are slightly unappreciated ,this study showed the potential of encoder-decoder deep learning models for such application.

10.2 Future Work

Although this project showed the potential of using deep learning for classifying cloud structures and gave less appreciated results , there were limitations encountered and ideas that could have been implemented to boost the project outcome. Therefore, this section provided recommendations for future work that can be implemented. As a future work , the models can be trained with the original images size and try different batch sizes. It also recommended to use the segmentation results and classify it further using classification models ,setting a probabilistic threshold , which ultimately results in a better classification, in another word , segment each region of the image belonging to a class and feed it further to a classfier. Additionally, the usage of more dataset will improve the model by having more samples to train the model with.

11 Student Reflections

There are many things, which I have not experienced before ,which I experienced in this project .One of it being the complexity and scale. However , it was a great opportunity for me to expand my skills and knowledge in the field of deep learning , image segmentation and programming skills , which is a transferable skill for my future career and study. Regardless of the challenges I faced , I was able to finish the project within the strict timeframe and that was achieved by following ,strictly the management plan I had created in the beginning of the project. The main problem I faced on this project is hardware , due to the long run of model to obtain results and limitation with GPU , nevertheless by putting more effort into the project and supervisor guidance I was able to get results. Finally , I could say lots of skills were gained throughout this project ,which I am confident to say it is golden and transferable.

Bibliography and References

[1] N.d, 2021. [online] Nasa.gov. Available at: <https://www.nasa.gov/pdf/135641main_clouds_trifold21.pdf> [Accessed 12 July 2021].

[2] Twomey, S. (1974). Pollution and the planetary albedo. *Atmospheric Environment (1967)*, *8*(12), 1251-1256. https://doi.org/10.1016/0004-6981(74)90004-3

[3] Dowling, D., & Radke, L. (1990). A Summary of the Physical Properties of Cirrus Clouds. *Journal Of Applied Meteorology*, *29*(9), 970-978. https://doi.org/10.1175/1520-0450(1990)029<0970:asotpp>2.0.co;2

[4] McLean, G. (1957). Cloud Distributions in the Vicinity of Jet Streams. *Bulletin Of The American Meteorological Society*, *38*(10), 579-583. https://doi.org/10.1175/1520-0477-38.10.579

[5] Rasp, S., Schulz, H., Bony, S., & Stevens, B. (2020). Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection. *Bulletin Of The American Meteorological Society*, *101*(11), E1980-E1995. https://doi.org/10.1175/bams-d-19-0324.1

[6] Shu, Y. (2014). *Deep Convolutional Neural Networks for Object Extraction from High Spatial Resolution Remotely Sensed Imagery* (PhD). University of Waterloo.

[7] Cai, S., & Liu, D. (2013). A comparison of object-based and contextual pixel-based classifications using high and medium spatial resolution images. *Remote Sensing Letters*, *4*(10), 998-1007. https://doi.org/10.1080/2150704x.2013.828180

[8] Rasp, S., Schulz, H., Bony, S., & Stevens, B. (2020). Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection. *Bulletin Of The American Meteorological Society*, *101*(11), E1980-E1995. https://doi.org/10.1175/bams-d-19-0324.1

[9] Turner, D., Vogelmann, A., Austin, R., Barnard, J., Cady-Pereira, K., & Chiu, J. et al. (2007). Thin Liquid Water Clouds: Their Importance and Our Challenge. *Bulletin Of The American Meteorological Society*, *88*(2), 177-190. https://doi.org/10.1175/bams-88-2-177

[10] Muhlbauer, A., McCoy, I., & Wood, R. (2014). Climatology of stratocumulus cloud morphologies: microphysical properties and radiative effects. *Atmospheric Chemistry And Physics*, *14*(13), 6695-6716. https://doi.org/10.5194/acp-14-6695-2014

[11] N.d. (2021). *Understanding Clouds from Satellite Images | Kaggle*. Kaggle.com. Retrieved 9 May 2021, from https://www.kaggle.com/c/understanding_cloud_organization.

[12] Rieck, M., Nuijens, L., & Stevens, B. (2012). Marine Boundary Layer Cloud Feedbacks in a Constant Relative Humidity Atmosphere. *Journal Of The Atmospheric Sciences*, *69*(8), 2538-2550. https://doi.org/10.1175/jas-d-11-0203.1

[13] Norris, J. (1998). Low Cloud Type over the Ocean from Surface Observations. Part II: Geographical and Seasonal Variations. *Journal Of Climate*, *11*(3), 383-403. https://doi.org/10.1175/1520-0442(1998)011<0383:lctoto>2.0.co;2

[14] Garay, M., Davies, R., Averill, C., & Westphal, J. (2004). ACTINOFORM CLOUDS: Overlooked Examples of Cloud Self-Organization at the Mesoscale. *Bulletin Of The American Meteorological Society*, *85*(10), 1585-1594. https://doi.org/10.1175/bams-85-10-1585

[15] Tapakis, R., & Charalambides, A. (2013). Equipment and methodologies for cloud detection and classification: A review. *Solar Energy*, *95*, 392-430. https://doi.org/10.1016/j.solener.2012.11.015

[16] Xie, W., Liu, D., Yang, M., Chen, S., Wang, B., & Wang, Z. et al. (2020). SegCloud: a novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation. *Atmospheric Measurement Techniques*, *13*(4), 1953-1961. https://doi.org/10.5194/amt-13-1953-2020

[17] Dev, S., Lee, Y., & Winkler, S. (2017). Color-Based Segmentation of Sky/Cloud Images From Ground-Based Cameras. *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, *10*(1), 231-242. https://doi.org/10.1109/jstars.2016.2558474

[18] Yuan, K., Meng, G., Cheng, D., Bai, J., Xiang, S., & Pan, C. (2017). Efficient cloud detection in remote sensing images using edge-aware segmentation network and easy-to-hard training strategy. *2017 IEEE International Conference On Image Processing (ICIP)*. https://doi.org/10.1109/icip.2017.8296243

[19] Shi, M., Xie, F., Zi, Y., & Yin, J. (2016). Cloud detection of remote sensing images by deep learning. *2016 IEEE International Geoscience And Remote Sensing Symposium (IGARSS)*. https://doi.org/10.1109/igarss.2016.7729176

[20] Azimi-Sadjadi, M., & Zekavat, S. (2000). Cloud classification using support vector machines. *IGARSS 2000. IEEE 2000 International Geoscience And Remote Sensing Symposium. Taking The Pulse Of The Planet: The Role Of Remote Sensing In Managing The Environment. Proceedings (Cat. No.00CH37120)*. https://doi.org/10.1109/igarss.2000.861666

[21] Qing Zhang, & Chunxia Xiao. (2014). Cloud Detection of RGB Color Aerial Photographs by Progressive Refinement Scheme. *IEEE Transactions On Geoscience And Remote Sensing*, *52*(11), 7264-7275. https://doi.org/10.1109/tgrs.2014.2310240

[22] Lee, J., Weger, R., Sengupta, S., & Welch, R. (1990). A neural network approach to cloud classification. *IEEE Transactions On Geoscience And Remote Sensing*, *28*(5), 846-855. https://doi.org/10.1109/36.58972

[23] Zhang, J., Liu, P., Zhang, F., & Song, Q. (2018). CloudNet: Ground-Based Cloud Classification With Deep Convolutional Neural Network. *Geophysical Research Letters*, *45*(16), 8665-8672. https://doi.org/10.1029/2018gl077787

[24] Jeppesen, J., Jacobsen, R., Inceoglu, F., & Toftegaard, T. (2019). A cloud detection algorithm for satellite imagery based on deep learning. *Remote Sensing Of Environment*, *229*, 247-259. https://doi.org/10.1016/j.rse.2019.03.039

[25] Zhai, H., Zhang, H., Zhang, L., & Li, P. (2018). Cloud/shadow detection based on spectral indices for multi/hyperspectral optical remote sensing imagery. *ISPRS Journal Of Photogrammetry And Remote Sensing*, *144*, 235-253. https://doi.org/10.1016/j.isprsjprs.2018.07.006

[26] Li, P., Dong, L., Xiao, H., & Xu, M. (2015). A cloud image detection method based on SVM vector machine. *Neurocomputing*, *169*, 34-42. https://doi.org/10.1016/j.neucom.2014.09.102

[27] Wang, S., Chen, W., Xie, S., Azzari, G., & Lobell, D. (2020). Weakly Supervised Deep Learning for Segmentation of Remote Sensing Imagery. *Remote Sensing*, *12*(2), 207. https://doi.org/10.3390/rs12020207

[28] Jin, W., Wang, L., Zeng, X., Liu, Z., & Fu, R. (2014). Classification of clouds in satellite imagery using over-complete dictionary via sparse representation. *Pattern Recognition Letters*, *49*, 193-200. https://doi.org/10.1016/j.patrec.2014.07.015

[29] N.d. (2018). Artificial Neural Network (ANN). Retrieved 13 July 2021, from http://www.cs.kumamoto-u.ac.jp/epslab/ICinPS/Lecture-2.pdf

[30] Zhu, X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience And Remote Sensing Magazine*, *5*(4), 8-36. https://doi.org/10.1109/mgrs.2017.2762307

[31] Muruganandham, S. (2016). *Semantic Segmentation of Satellite Images using Deep Learning* (MSc). Luleå University of Technology.

[32] Zhu, X., Tuia, D., Mou, L., Xia, G., Zhang, L., Xu, F., & Fraundorfer, F. (2017). Deep Learning in Remote Sensing: A Comprehensive Review and List of Resources. *IEEE Geoscience And Remote Sensing Magazine*, *5*(4), 8-36. https://doi.org/10.1109/mgrs.2017.2762307

[33] Ghosh, A., Sufian, A., Sultana, F., Chakrabarti, A., & De, D. (2019). Fundamental Concepts of Convolutional Neural Network. *Intelligent Systems Reference Library*, 519-567. https://doi.org/10.1007/978-3-030-32644-9_36

[34] Noh, H., Hong, S., & Han, B. (2015). Learning Deconvolution Network for Semantic Segmentation. *2015 IEEE International Conference On Computer Vision (ICCV)*. https://doi.org/10.1109/iccv.2015.178

[35] N.d. (2021). *What Is Agile Methodology in Project Management?*. Wrike.com. Retrieved 10 July 2021, from https://www.wrike.com/project-management-guide/faq/what-is-agile-methodology-in-project-management/.

[36] Muaz, U. (2020). *Rethinking Data Augmentations: A causal perspective*. Medium. Retrieved 4 July 2021, from https://towardsdatascience.com/tagged/data-augmentation?p=e0b7810579a7.

[37] Tan, M., & Le, Q. (2020). *EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks*. arXiv.org. Retrieved 1 July 2021, from https://arxiv.org/abs/1905.11946v3.

[38] Lee, J., Won, T., Lee, T., Lee, H., Gu, G., & Hong, K. (2020). *Compounding the Performance Improvements of Assembled Techniques in a Convolutional Neural Network*. arXiv.org. Retrieved 1 August 2021, from https://arxiv.org/abs/2001.06268.

[39] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *2018 IEEE/CVF Conference On Computer Vision And Pattern Recognition*. https://doi.org/10.1109/cvpr.2018.00474

[40] Ronneberger, O., Fischer, P., & Brox, T. (2015). *U-Net: Convolutional Networks for Biomedical Image Segmentation*. arXiv.org. Retrieved 1 July 2021, from https://arxiv.org/abs/1505.04597v1.

[41] Oktay, O., Schlemper, J., Folgoc, L., Lee, M., Heinrich, M., & Misawa, K. et al. (2018). *Attention U-Net: Learning Where to Look for the Pancreas*. arXiv.org. Retrieved 2 July 2021, from https://arxiv.org/abs/1804.03999v2.

[42] Chen, X., Yao, L., & Zhang, Y. (2020). *Residual Attention U-Net for Automated Multi-Class Segmentation of COVID-19 Chest CT Images*. arXiv.org. Retrieved 3 July 2021, from https://arxiv.org/abs/2004.05645.

[43] He, K., & Sun, J. (2015). Convolutional neural networks at constrained time cost. *2015 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2015.7299173

[44] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. *2016 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2016.90

[45] Xie, S., Girshick, R., Dollar, P., Tu, Z., & He, K. (2017). Aggregated Residual Transformations for Deep Neural Networks. *2017 IEEE Conference On Computer Vision And Pattern Recognition (CVPR)*. https://doi.org/10.1109/cvpr.2017.634

[46] N.d. (2021). TensorFlow. TensorFlow. Retrieved 7 July 2021, from https://www.tensorflow.org/.

[47] Zaccone, G., Karim, M., & Menshawy, A. (2017). Deep Learning with TensorFlow (pp. 10-113). Packt Publishing.

[48] Lee, A. (2019). Announcement: TensorFlow 2.0 Has Arrived!. Medium. Retrieved 7 July 2021, from https://towardsdatascience.com/announcement-tensorflow-2-0-has-arrived-ee59283fd83a.

[49] Liu, L., Jiang, H., He, P., Chen, W., Liu, X., Gao, J., & Han, J. (2020). On the Variance of the Adaptive Learning Rate and Beyond. arXiv.org. Retrieved 7 July 2021, from https://arxiv.org/abs/1908.03265.

[50] Jadon, S. (2020). A survey of loss functions for semantic segmentation. Retrieved 7 July 2021, from https://arxiv.org/pdf/2006.14822.pdf.

[51] Sudre, C., Li, W., Vercauteren, T., Ourselin, S., & Jorge Cardoso, M. (2017). Generalised Dice Overlap as a Deep Learning Loss Function for Highly Unbalanced Segmentations. Retrieved 7 July 2021, from https://arxiv.org/abs/1707.03237.

[52] Dice, L. (1945). Measures of the amount of ecologic association between species. *Ecology*, *26*(3), 297–302

[53] Berman, M., Triki, A., & Blaschko, M. (2018). *The Lov\'asz-Softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks*. arXiv.org. Retrieved 11 July 2021, from https://arxiv.org/abs/1705.08790. (Berman et al., 2018)

Appendix A – Meeting Records

Meeting No	Discussed Topics	Data	Time
1	• On this meeting , different possible projects were discussed to be considered as the final project such as satellite image segmentation and automatic number plate detection. • Supervisor suggested websites as potential source for satellite images data	06/04/2021	14:30-15:00
2	• It was suggested by supervisor to explore image data from NOAA website for the purpose of feature extraction. • Consider the option of using webcam to create data.	20/04/2021	14:30-15:00
3	• Group project meeting , discussed all aspect of the project such as the process of the project , time constrains and expected outcome.	15/05/2021	13:00-13:300
4	• It was discussed to explore various options to gather data as well as exploring feature extraction techniques such as neural network. • It was also discussed to initiate literature review to have better understanding of the project.	28/05/2021	13:00-13:300
5	• Discussed the project proposal and submit it in the allocated deadline. • It was suggested to continue working on the literature review as well as training and testing the models.	11/06/2021	13:00-13:300
6	• Similarly, with previous meeting after discussing the results obtained , it was suggested to continue training the models and evaluate the results. • Keep working on the literature review. • It was also suggested by supervisor to resize the images for easy processing to overcome computing problems .	25/06/2021	16:30-17:00
7	• Continue making progress with development of the models. • Discussed the deadlines of the presentation and final report.	12/07/2021	16:00-16:30
8	• Finalised the project results . • It was advised by supervisor to pace work to meet deadlines.	26/07/2021	16:00-16:30

Appendix B – Project Presentation

Understanding and Classifying Clouds Structure In Satellite Images Using Supervised Deep Learning Techniques

Name: Abas Mohammedidris (SID: 8931834)

Course: Data Science & Computational Intelligence

Supervisor: Dr. Aamer Saleem

Second Supervisor: Dr. Diana Nkantah

Academic Year: 2020/21

Coventry University

Introduction

Problem Statement
- Literature Review
- Dataset
- Methodology
- Research Results

Discussion

Future Work

Conclusion / Findings

Presentation Outline

Introduction

01

- One of the most discussed topics in the forefront in recent years is climate change.
- Cloud play an important role in climate change and that's by controlling the amount of solar energy that reaches the planet surface and the amount of the planet's energy which radiated back to the space.
- Cargo ships and aircraft flights ,carry more than 70% of the world trade shipment and millions of passengers and are highly dependent on the weather.

02

- To interpret cloud's structure in satellite images ,human being expertise is used.
- In recent years, many approaches were proposed to build an automated and intelligent system to interpret cloud's structure.

03

- Cloud's structure interpretation using deep learning require a big amount of labelled data.
- Having a good dataset is crucial for further study.
- Detecting clouds formation beforehand can help to avoid the loss of lives and millions of dollars.

04

- This study aims to answer the following question: ***Can Deep Learning be used to interpret cloud's structure and classify them?***
- The findings of this study will aid climatologist , cargo ships and aviation industry.

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Problem Statement

1. Literature Review

- Analyze current state-of-the-art techniques
- Examine previous related work

2. Dataset

- Data Presentation
- Data Investigation
- Data Pre - Processing

3. Methodology

- Introduce Semantic Segmentation models for our dataset.
- Develop models for segmenting clouds structures in satellite images.

4. Research Results

- Examine the performance of the models
- Present visual results and tabulate results.

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Literature Review

- (Li et al., 2015) Used SVM for cloud detection and the method gave an accuracy of 90% for clouds detection.
- (Xie et al., 2020) Introduced a technique named SegCloud for cloud segmentation and the model consist of encoder-decoder path, this method achieved a promising results in clouds segmentation.
- (Dev et al., 2017) proposed a colour-based cloud image segmentation and this technique was tested on Singapore whole sky dataset.
- (Jeppesen et al., 2019) Used a U-Net architecture-based model for detecting clouds in different regions such as regions with snow and gave a promising.

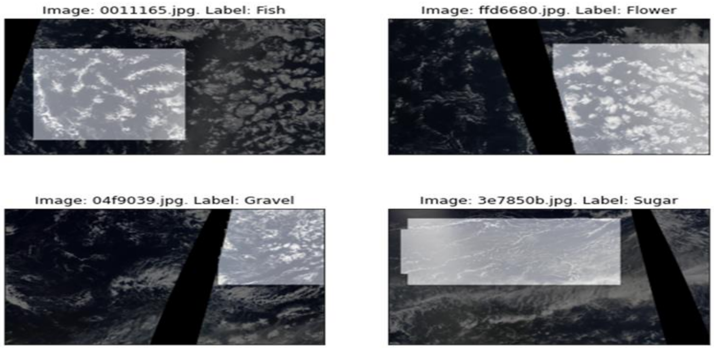
Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Data Presentation

- The original dataset was collected by by NASA worldwide view and it was taken from three regions spanning in 21° longitude and 14° latitude, the original-colour images were collected from Terra and Aqua , whose are two-orbiting satellites.
- The uncovered area by the two succeeding orbits is marked in black .
- The dataset consist of two files , training, and testing. The training dataset consist of 5546 × 4 and the testing dataset has 3698 images.
- The training set was labelled by 68 different scientist and each image was labelled by three different individuals(Rasp et al., 2020) .
- The dataset was uploaded into Kaggle competition by (N.d,2019)
- The Dataset consist of four clouds types namely fish , flower , sugar and gravel.

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

- Sugar: describes in areas of very fine cumulus in widespread manner , this class of clouds exhibits the mesoscale organization ,because the overall field does not have much of cloud-free region.
- Flower : the areas described by flower are the clouds that have isotropic structures , which have a range of 50km to 200 km in diameter , which consist of similar cloud-free region in between and its is densely packed.
- Fish :the fish feature or patterns are elongated ,which sometime spin up to 1000 km , in most cases in longitudinal manner, in previous literature this pattern was called action-from clouds.
- Gravel: this pattern name describes the fields that have granular features whose are marked by rings. The normal scale of these rings is approximately 20 km , it is highly suspected that ,this pattern is driven by cold pools thar are caused by raining cumulus.

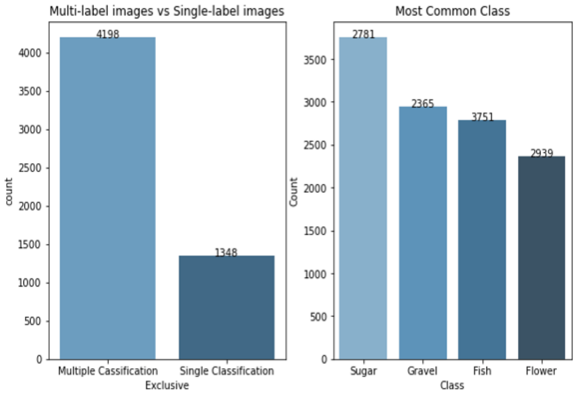


Four clouds types

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Data Investigation

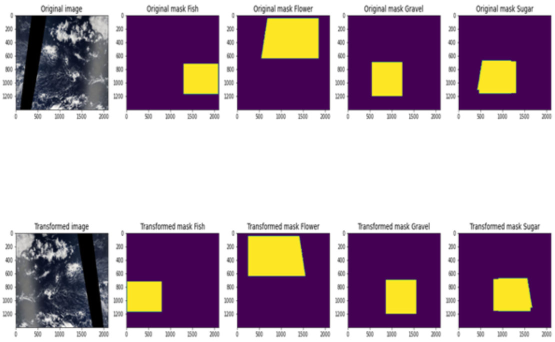
- The dataset is multi-class and multi-label , which means an image can have a number of labels between one and four.
- The most occurrence class is sugar, and the least is flower.



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Data Pre - Processing

- **The original images size was 2100x1400x3, but it was resized to 128x256x3 due to computing power limitations.**
- **To increase the training dataset ,data augmentation was used to have more data for training the models and improve the performance .**
- **The images have been augmented into five types vertical file , horizontal flip , random rotation 90° , grid distortion and optical distortion**
- **The dataset was split 80:20 into training and validation.**



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

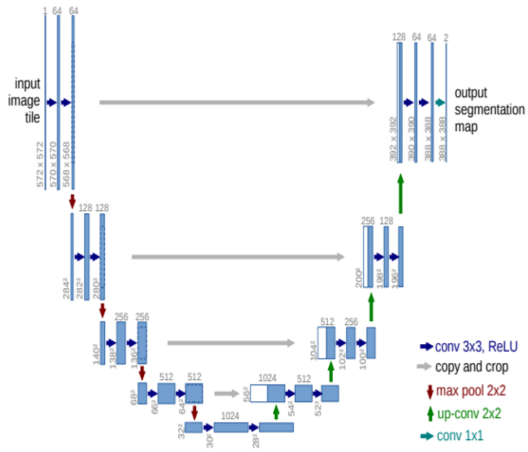
Methodology

Phase 1	Phase 2	Phase 3	Phase 4	Phase 5
Models Investigation	**Models Design**	**Implementation**	**Models Evaluation**	**Prediction**
• U-Net • Attention-UNet • Residual Attention_Unet • EfficientNetB7_Unet • EfficientNetB6_Unet	• Optimizers			

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

U-Net

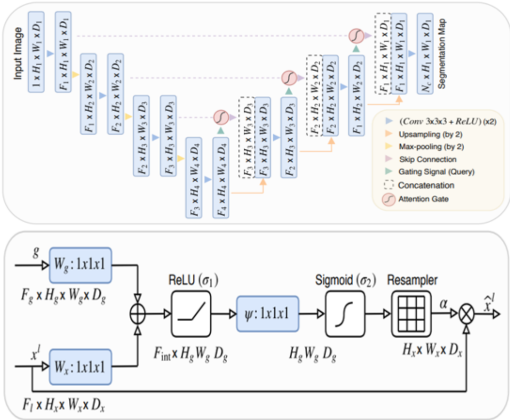
- **The U-Net consists of two paths ,contracting path and expansion path.**
- **The contracting path is normal convolutional network ,that have repeated convolutions,where each convolution is followed by an activation function (ReLU) and max pooling, spatial information is reduced, and feature maps are increased in this path.**
- **The second path uses, up-convolutions and concatenations to combine the feature and spatial information and give an output with high resolution.**



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Attention-UNet

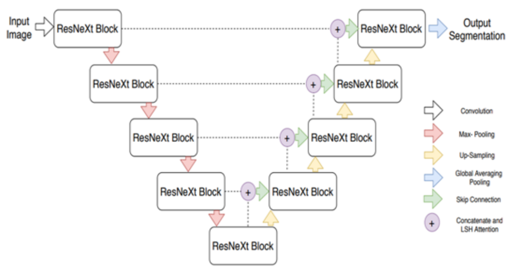
- The Attention-Unet have two paths like U-Net , the only difference is ,Attention_Unet uses attention gates to highlight more important features.
- The AG takes two inputs ,namely vector “X” and “g” , where “X” comes from the skip connection, and it has better spatial information and “g” comes from the next lowest layer of the network and since it comes from deeper part of the network it has better feature representation.
- the AGs filters the neurons as well as the irrelevant information from the background are downweighed.
- This allows the layers only to be updated with most relevant regions based on the task.



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Residual Attention_Unet

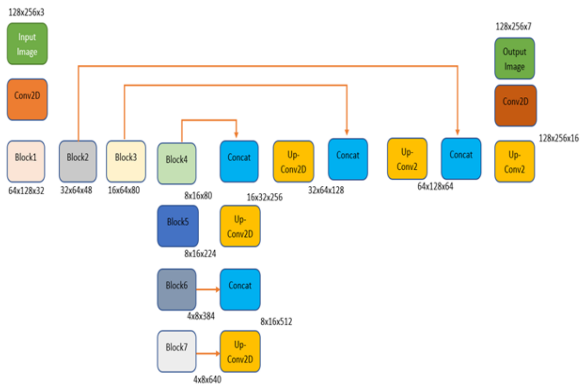
- The Residual Attention_Unet has similar architecture ,the only difference is that ResNeXt blocks are featured within Residual Attention U-Net to capture features from the original input image .
- Aggregated Residual Network (ResNeXt) increases the cardinality instead of depth .



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

EfficientNetB7_Unet

- Like the previous networks , this network has contracting and expansion path.
- In this network , EfficientNetB6 and EfficientNetB7 are used as encoders or contracting path.
- The EfficientNet consist of 8 models namely B0-B8 and each model differ from another by number of parameters and performance.
- The EfficientNet operates in three ways , depth wise + pointwise convolution , Inverse Residual and bottleneck where it applies linear activation .



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Model Design and Implementation

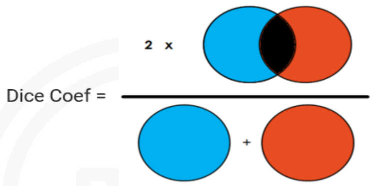
- All the models were implemented in TensorFlow 2
- Rectified Adam optimiser was used to optimise the learning rate .
- Early stop was used , to stop the model training if the model validation loss do not improve.
- The activation function selected to be used in the layers was `relu' and for the output layer `softmax' was employed.
- All the models were regularized except the EfficientUNet models.

Model	Total parameters	Trainable parameters	Non trainable
U-Net	31,402,708	31,390,924	11,784
Attention U-Net	37,334,872	37,319,248	15,624
Residual Attention U-Net	39,090,584	39,069,072	21,512
EfficientNetB6-UNet	37,435,258	37,227,151	208,107
EfficientNetB7-UNet	53,564,192	53,288,201	275,991

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

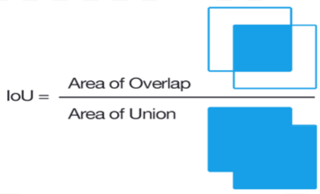
Model Evaluation

Dice Coefficient



- Dice coefficient is one of most used metrice for model evaluation for semantic segmentation. Dice Coefficient is simply , 2 multiplied by the area of overlap divided by the area of both images.

Intersection over Union



- Another technique used to evaluate models ,when it comes to semantic segmentation is intersection over union(IoU) or jacard index ,which is simply the area overlap divided by the area of union.
- Both evaluation metrics produce similar evaluation results ,if Dice coefficient says model A is good ,IoU will say similar thing.
- Both range between 0 and 1 ,with 1 signifying the highest similarity between prediction and ground truth.

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

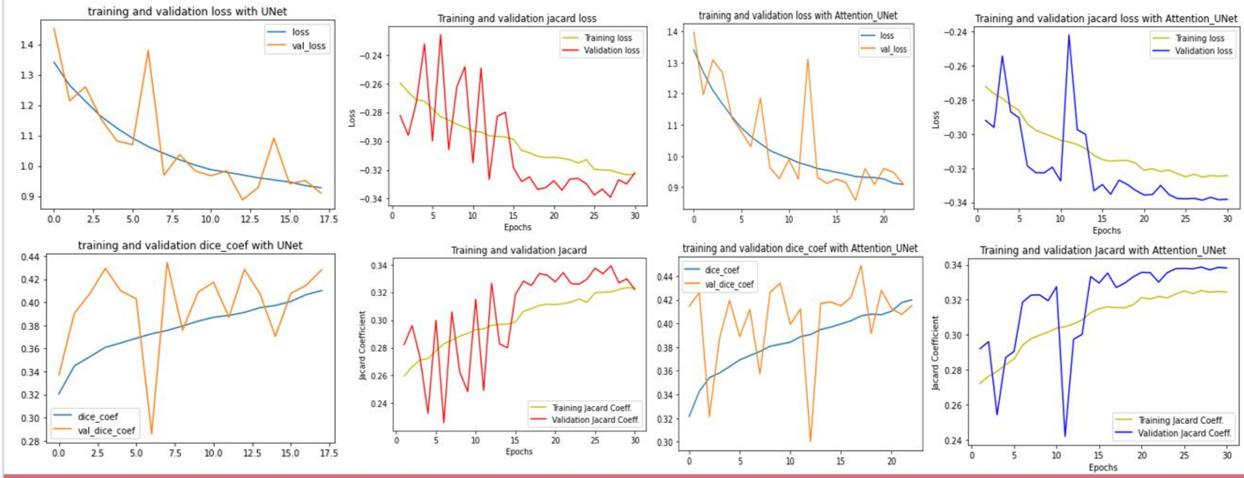
Research Results and Discussion

- From the evaluation results ,we observe,EfficientNetB6-UNet has the best performance ,with dice coefficient of 0.53 on the validation and 0.59 on the training.
- The second best model is EfficientNetB7-Unet.
- The models using EfficientNet as encoder have the best performance among the models.

Model	Dice Coefficient		Intersection over Union	
	Training	Validation	Training	Validation
U-Net	0.4301	0.4282	0.3233	0.3221
Attention U-Net	0.4463	0.4488	0.3382	0.3243
Residual Attention U-Net	0.4600	0.4667	0.3403	0.3412
EfficientNetB6-UNet	0.5934	0.5329	0.4587	0.4045
EfficientNetB7-UNet	0.5503	0.5014	0.4531	0.4083

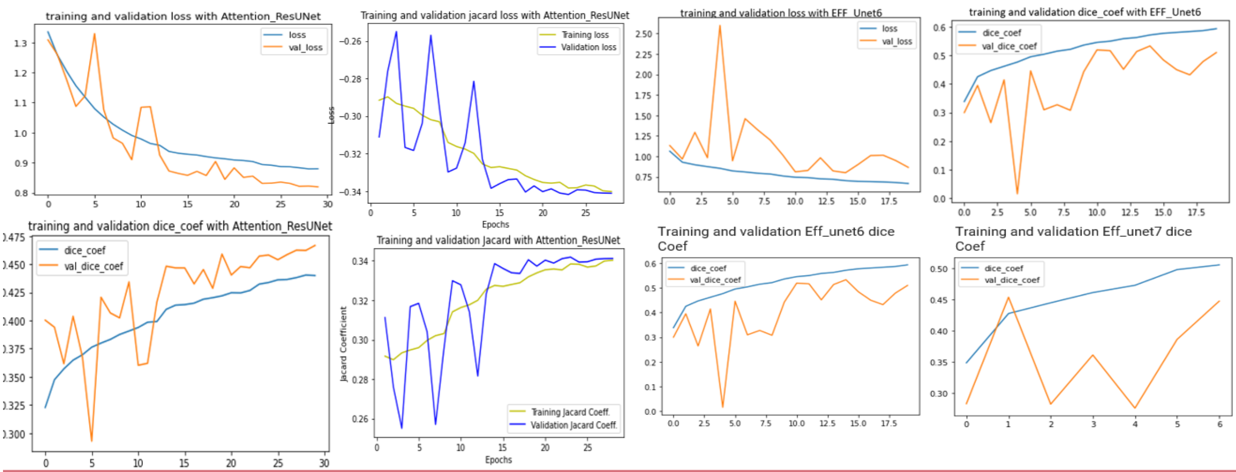
Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Results



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

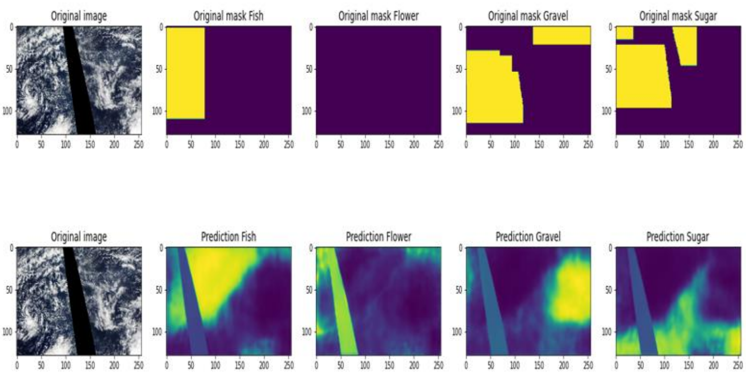
Results



Abas Mohammedidris Classification of Clouds Structure in satellite Images Using Deep Learning

Segmentation Results

- The models were tested on predicting masks and based on the results we observe the segmentation results presented in this slide .
- As we observe ,from the original image ,the similarity is very high between the classes ,in terms of patterns.



Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Discussion

Discussion Topic #1

A different models were deployed for the purpose of classifying clouds into four classes using segmentation models. and all models were evaluated using the same evaluation metrics.

Discussion Topic #2

All models were trained for 30 epochs ,however ,early stopping was used to stop the model if there's no improvement.

Discussion Topic #3

Th models using EfficientNet as encoder achieved good results on this dataset. With EfficientNetB6-Unet achieving a dice coefficient of 0.53 in validation data.

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Future Work

Suggestion #1

Use gradient weighted class activation mapping to generate a baseline, which is a class explainability technique.

Suggestion #2

- Train the models with the original image size, use more powerful computing and try different batches size.
- Feed the segmentation results into a classifier.

Suggestion #3

Replace the encoder path EfficientNet by more advanced models .

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

Conclusions / Findings

- This paper aims to make contributions to develop an automated intelligent system that interprets clouds patterns.
- Developing deep learning models that detects clouds and predict its formation patterns can help to avoid natural disasters such as hurricane and save lives as well as millions of dollars.
- In this study five different model ,were developed ,implemented and evaluated.
- This study showed by applying segmentation models for classification task , by using good encoder such as EfficientNet alongside U-Net ,good results can be attained.

Abas Mohammedidris Classification of Cloud Structures in Using deep learning

References

1 [Wang, S., Chen, W., Xie, S., Azzari, G., & Lobell, D. (2020). Weakly Supervised Deep Learning for Segmentation of Remote Sensing Imagery. Remote Sensing, 12(2), 207. https://doi.org/10.3390/rs12020207

2 [Xie, W., Liu, D., Yang, M., Chen, S., Wang, B., & Wang, Z. et al. (2020). SegCloud: a novel cloud image segmentation model using a deep convolutional neural network for ground-based all-sky-view camera observation. Atmospheric Measurement Techniques, 13(4), 1953-1961. https://doi.org/10.5194/amt-13-1953-2020

3 Rasp, S., Schulz, H., Bony, S., & Stevens, B. (2020). Combining Crowdsourcing and Deep Learning to Explore the Mesoscale Organization of Shallow Convection. Bulletin Of The American Meteorological Society, 101(11), E1980-E1995. https://doi.org/10.1175/bams-d-19-0324.1

5 [N.d. (2019). Understanding Clouds from Satellite Images | Kaggle. Kaggle.com. Retrieved 9 May 2021, from https://www.kaggle.com/c/understanding_cloud_organization .

6 [Dev, S., Lee, Y., & Winkler, S. (2017). Color-Based Segmentation of Sky/Cloud Images From Ground-Based Cameras. IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing, 10(1), 231-242. https://doi.org/10.1109/jstars.2016.2558474 (Dev et al., 2017)

THANK YOU !

Understanding and Classifying Clouds Structure In Satellite Images Using Supervised Deep Learning Techniques

Name: Abas Mohammedidris (SID: 8931834)

Course: Data Science & Computational Intelligence

Supervisor: Dr. Aamer Saleem

Second Supervisor: Dr. Diana Nkantah

Academic Year: 2020/21

QUESTIONS

Appendix C – Certificate of Ethics Approval

Certificate of Ethical Approval

Applicant:	Abas Mohammedidris
Project Title:	Understanding and Classifying Cloud Structures In Satellite Images Using Supervised and Semi-supervised Methods

This is to certify that the above named applicant has completed the Coventry University Ethical Approval process and their project has been confirmed and approved as Low Risk

Date of approval:	11 Jun 2021
Project Reference Number:	P122776

Figure 11.1: Certificate of ethical approval for the project

Appendix D – Augmentation Results

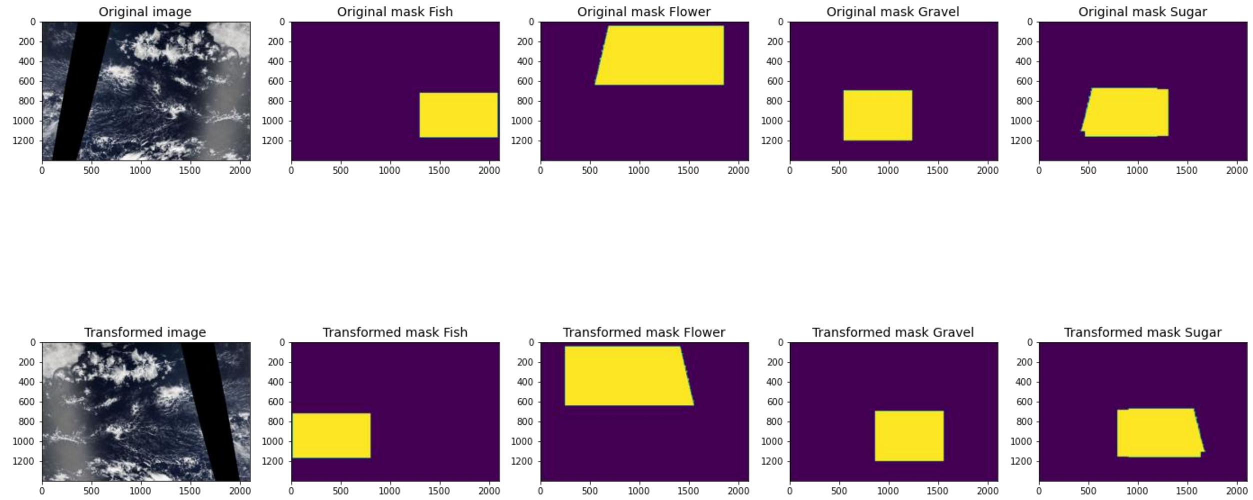


Figure 11.2: Augmentation results with Horizontal Flip

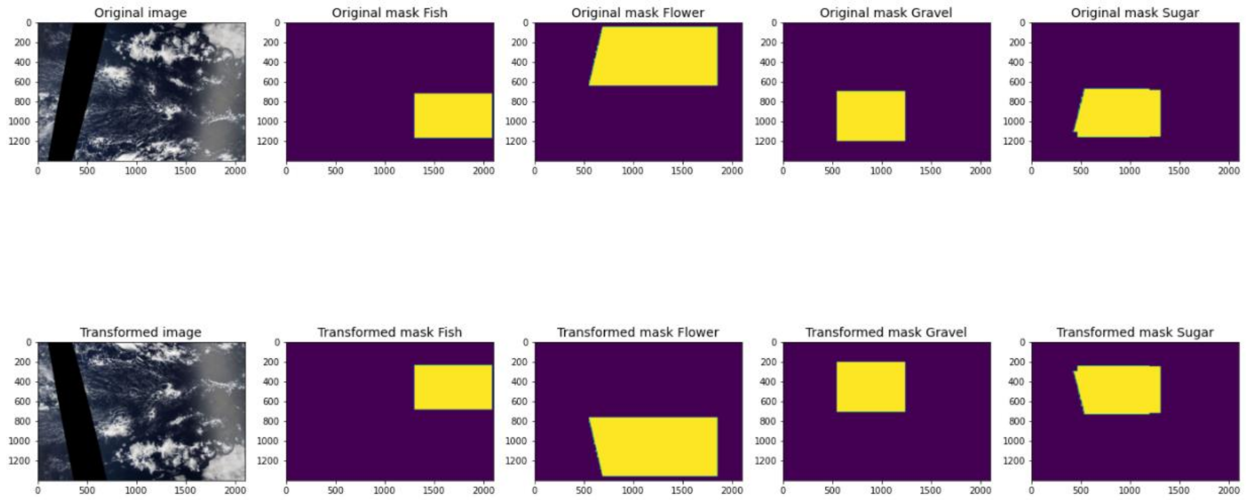


Figure 11.3: Augmentation results with vertical Flip

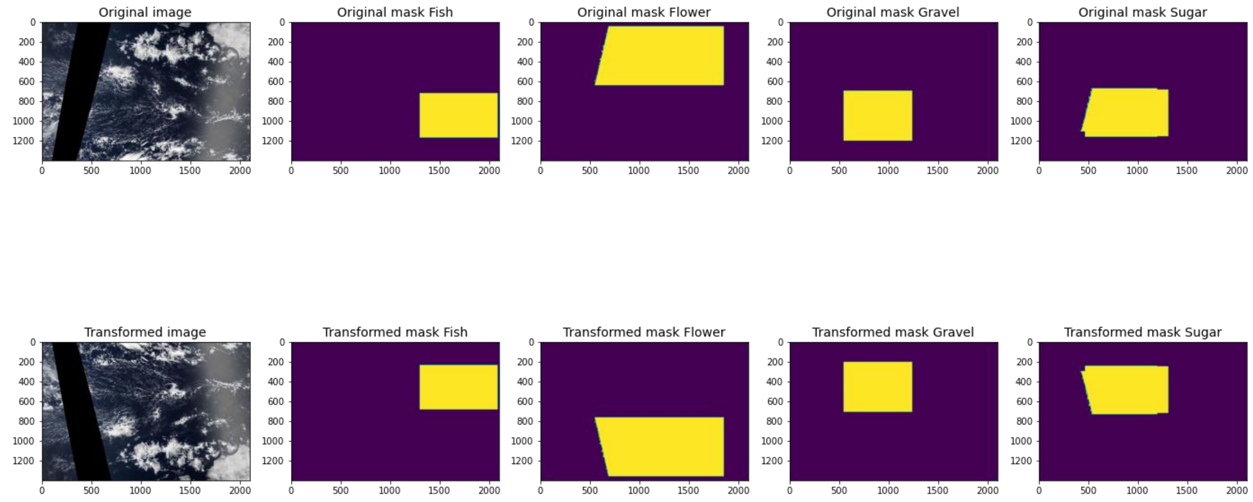


Figure 11.4: Augmentation results with RandomRotate90

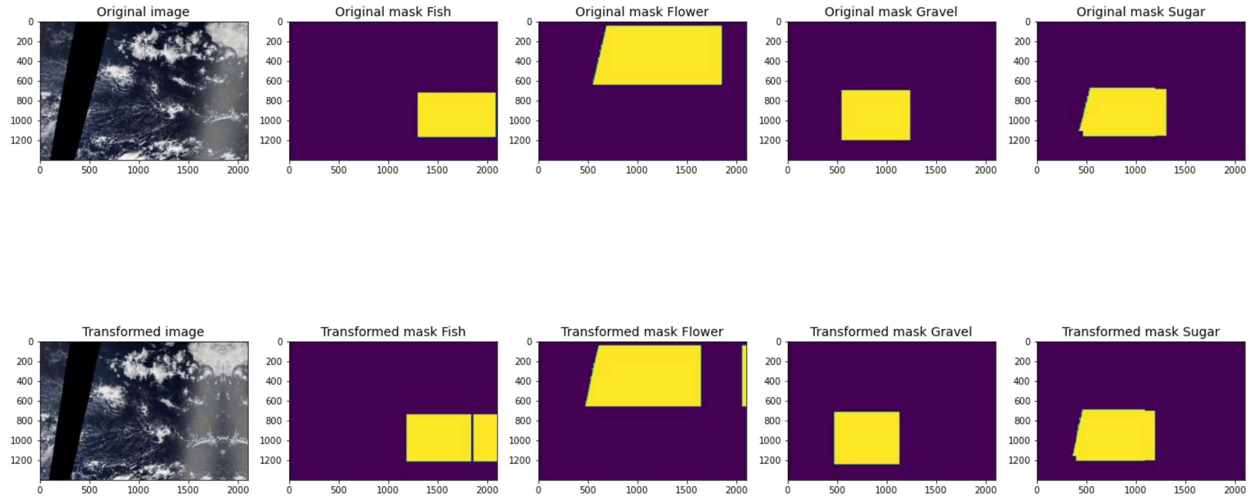


Figure 11.5:Augmentation results with GridDistortion

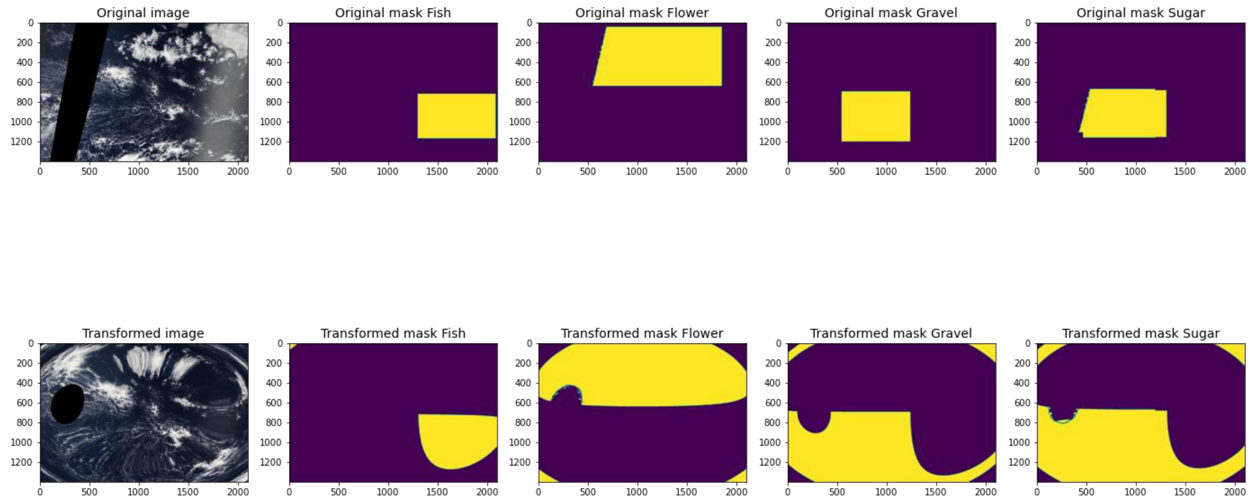


Figure 11.6:Augmentation results with OpticalDistortion

Appendix E – Segmentation Results

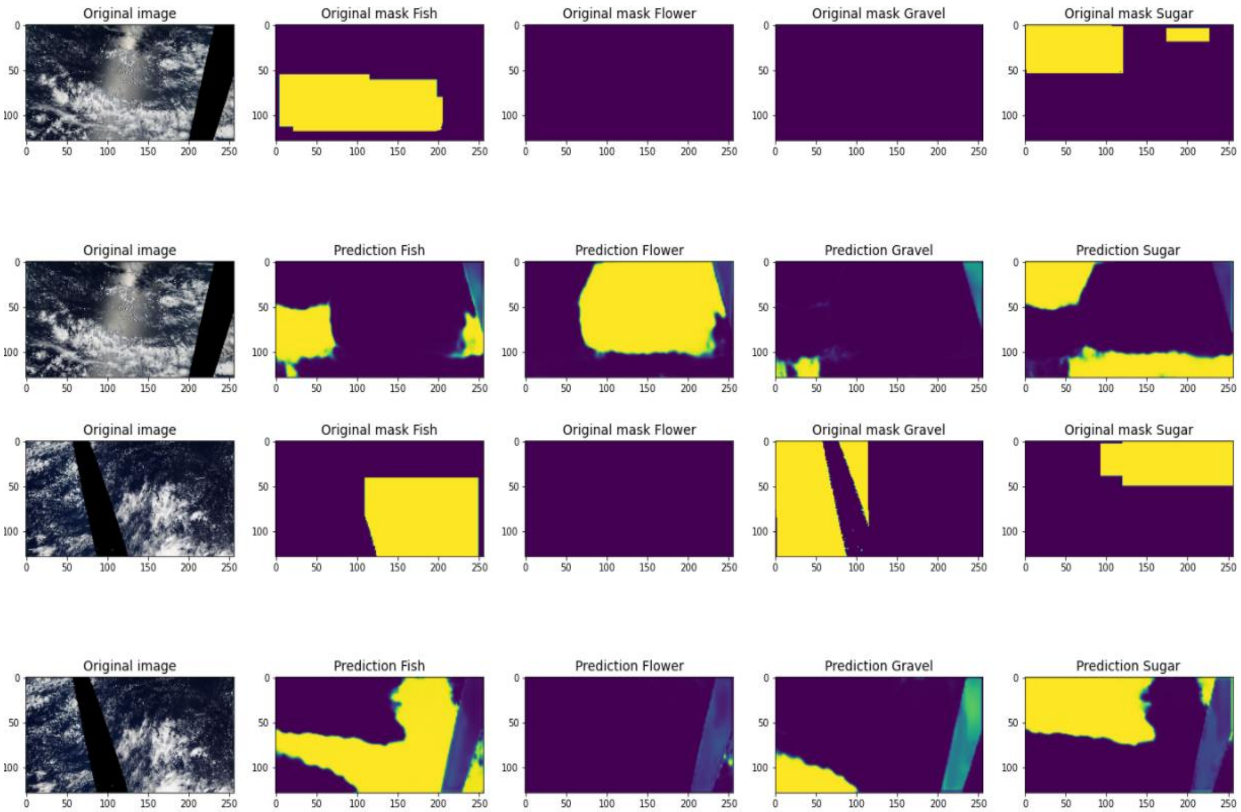


Figure 11.7:Sample of segmentation result with U-Net

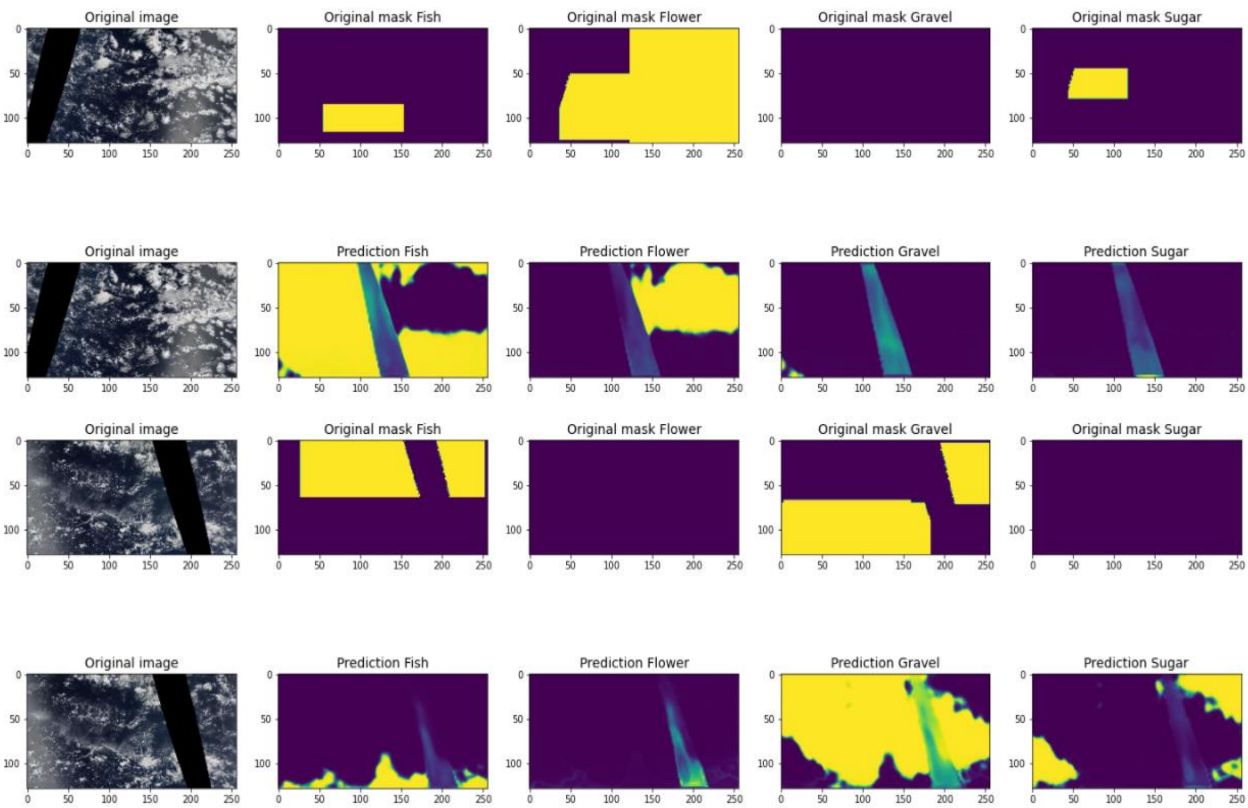


Figure 11.8:Sample of segmentation result with Attention U-Net

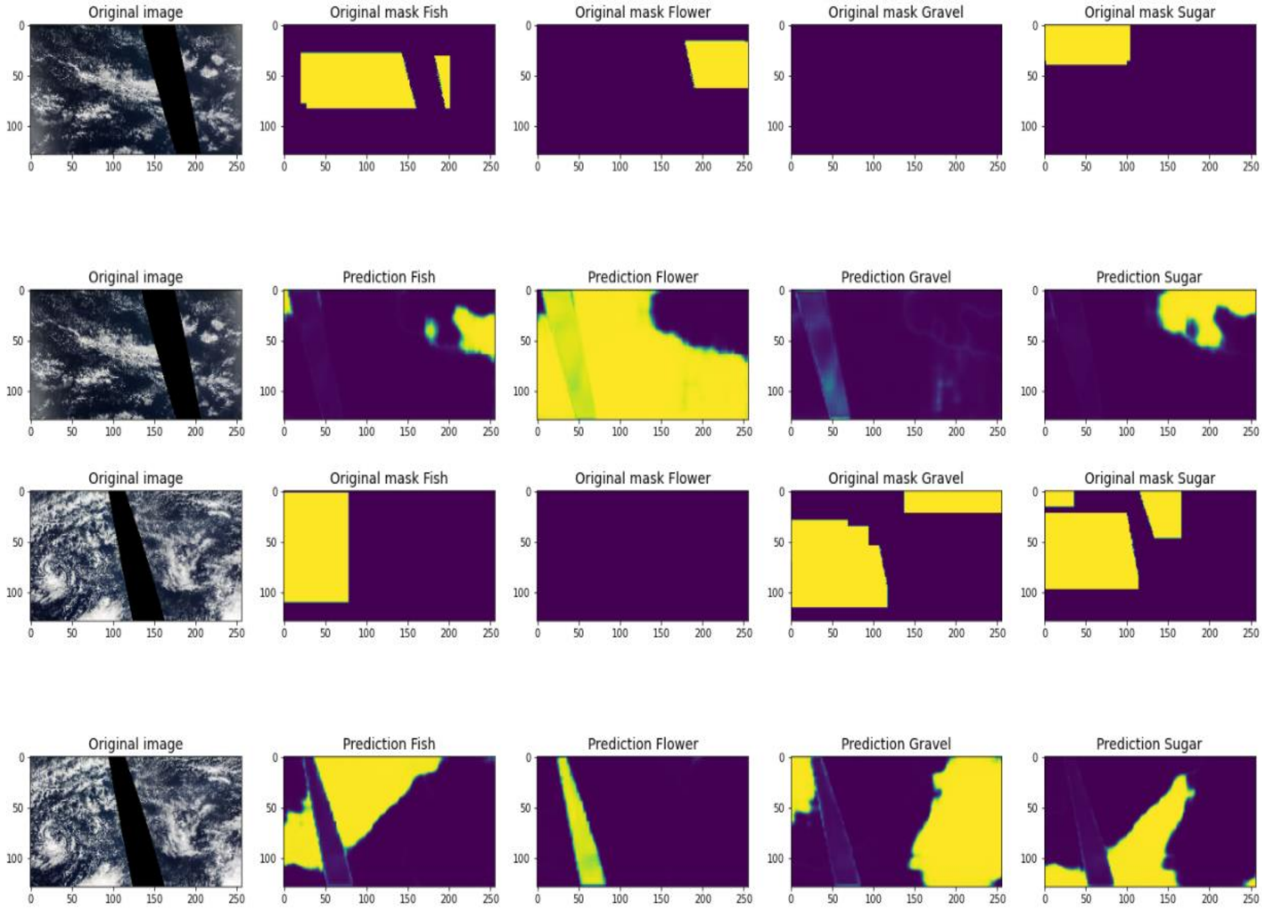


Figure 11.9: Sample of segmentation result with Residual Attention U-Net

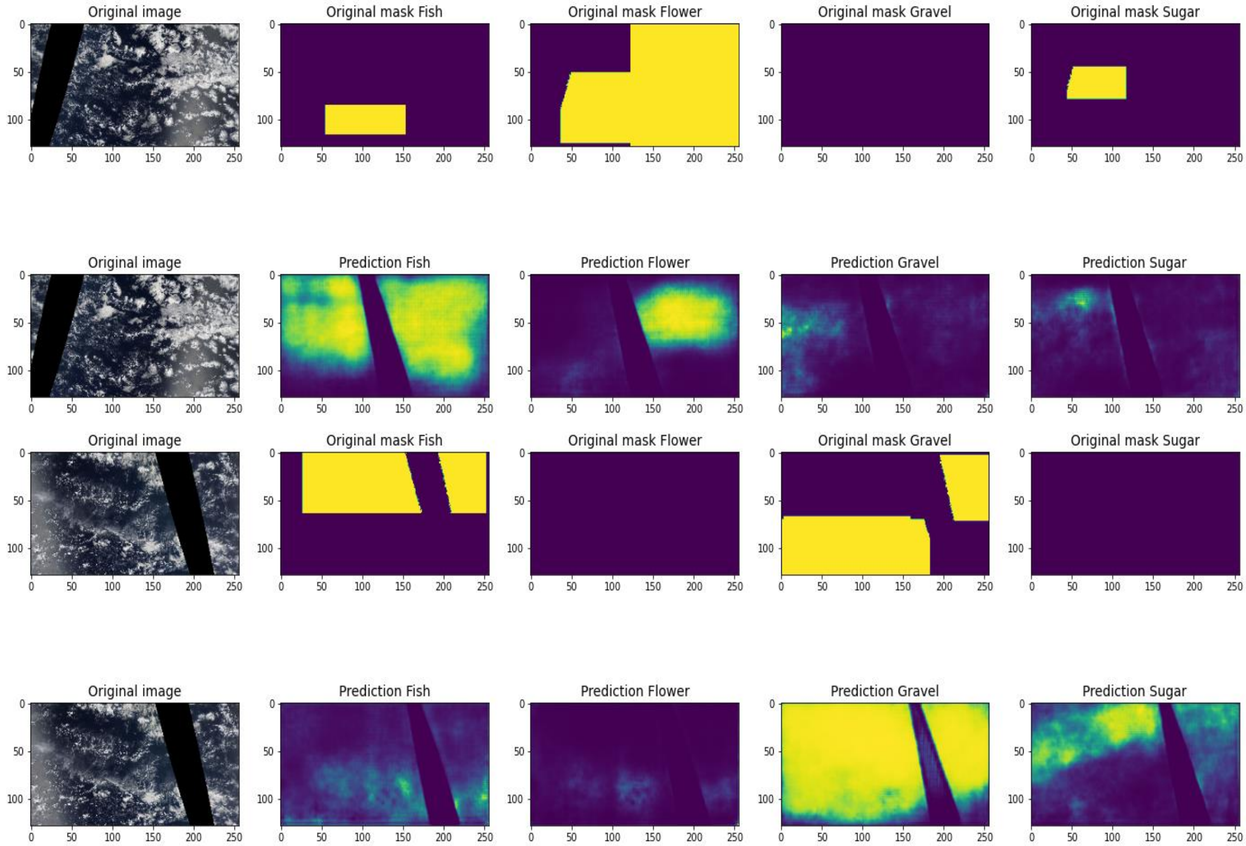


Figure 11.10: Sample of segmentation result with EfficientNetB6-Unet

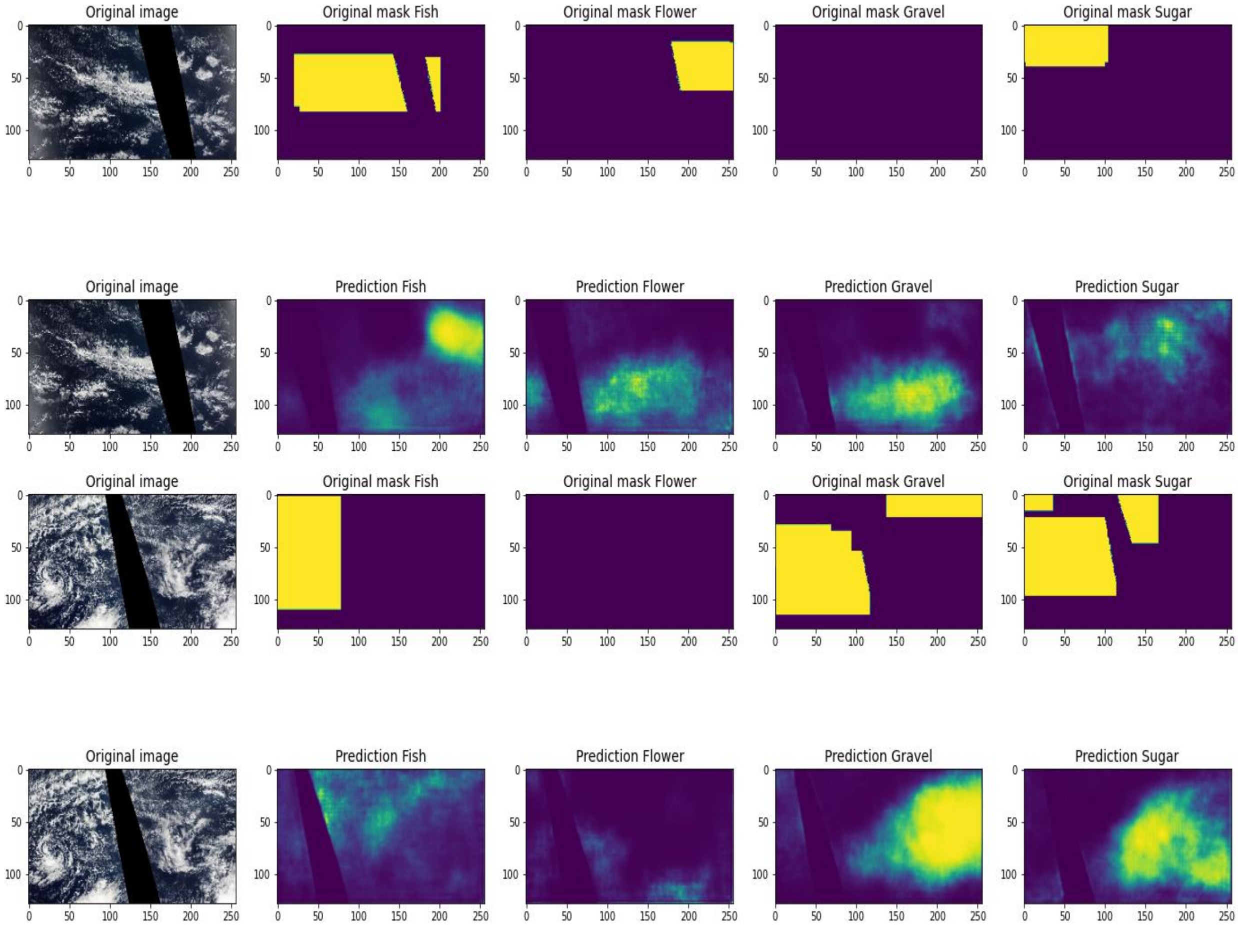


Figure 11.11:Sample of segmentation result with EfficientNetB7-Unet

Appendix F – Source code and Dataset

The source code and dataset of this project can be found in the links below. The source code is python ,it was also downloaded from jupyter notebook as html file, and it is included separately in the same folder.

The files are saved in two different locations ,in case one link does not work.

One drive link:

https://livecoventryac-my.sharepoint.com/:f:/g/personal/moham911_uni_coventry_ac_uk/ErmnWqE5DphDs92qA7JvDFoBcVHM9kXaU_yKq3rp76UxjQ?e=M8OOgB

google drive:

https://drive.google.com/drive/u/0/folders/1GFet8HA-cVy2gMf5xBgtGUm1jUt_9Q0t