Assignment 2, Cloud Pattern Classification using Satellite Images,2022
Due 11:59, 28/10/22

Priyank Tusharbhai Dave
The University of Adelaide
a1843068@adelaide.edu.in

Abstract— For many years, important political debates and decisions have focused heavily on the issue of climate change. Although they are difficult to comprehend and reflect in a climate model, shallow clouds are crucial to understanding the Earth's climate. By categorizing these cloud patterns, it is possible to better understand their physical makeup, which would enhance the creation of climate models and enable better predictions of climate change or weather forecasts [3]. Cloud organizations may take many different shapes, making it difficult to extract cloud elements using conventional rule-based methods. In this project, the classification of cloud patterns were carried out using a new augmented version and was used as a classifier, which will aid adroit in understanding how the cloud formation going to take place. It was demonstrated that utilizing a categorizing model in a classification job together with a strong encoder allowed UNet to perform well on this dataset.[1] The final assessment metric, the dice coefficient, yielded scores of 63.66% and 65.02% for the internal and external (test set) leaderboards in the Kaggle.[1]

1. Introduction

Clouds play a crucial part in regulating the sun's energy as well as the radiation that returns to the atmosphere. The planet's atmosphere warms up as more energy is held inside it, raising sea levels through the melting of glaciers and causing global warming. Less power that is contained, the cooler the environment gets. Understanding the context of clouds aids in a better comprehension of the planet's weather. As a result, climate researchers rely on it. Albedo is a parameter that expresses how much energy is radiated rather than absorbed.[4]
Dark surfaces absorb a large portion of energy, indicating a low albedo, whereas white surfaces emit a large amount of energy, showing a high albedo. The albedo of Earth is 0.3, which indicates that the temperature is warming [3]. Understanding how cloud shapes are interpreted can help us understand how the Earth's climate is being abused and the hazards that come with it. Clouds in satellite photographs provide a more comprehensive view of the environment, and their interpretation can bring information about the state of the globe right now.[4]

It is expected that when the global temperature rises, more water would evaporate from the seas, resulting in too many clouds with diverse forms and patterns. Climate change may be influenced by the overall effect of cloud forms and other characteristics such as abundance and thickness.[3]

Fig. 1. Four separate class samples or slimmer and altitude places were masked. Cloud constructions come in a variety of shapes and sizes. Cirrus clouds are the most common type of higher end cloud which means ‘Curly hair. These icicle-based, fluffy clouds have long, thin bands that are also referred to as mare's tails. Cirrus clouds that are dispersed signal pleasant weather. The advancement of a cirrus cloud that resembles a web signifies increased moisture in the air and the advent of a probable thunderstorm. Cirrostratus clouds spread their thin, dispersed lines over the sky, giving the air a subdued white appearance.[1]

These are some usual indications of impending rain. Cirrostratus cloud visibility is generally verified twelve to twenty-four hours before the advent of any storm. Another form of large cloud called cirrocumulus typically looks like a collection of neatly matched white lines. These might indicate the approach of a storm in tropical environments. Lateral cirrus bands may cause instability in planes, and heat maps indicate the geographical trends where the groups originate on a yearly basis.[1]

Many high-risk actions can be avoided in the past courtesy to cloud formation and the early detection of their shapes and patterns. Aviation is a hazardous activity that transports hundreds of people, So, knowing about cloud formations is a significant requirement for becoming a pilot. The lift force on wings is what induces turbulence, and cloud-borne updrafts and downdrafts have an immediate and unpredictable impact on this force. Turbulence is the term for these changes that make an airplane lurch and leap while it is in the air. Turbulence may often lead inexperienced pilots to lose control of their aircraft, which is frequently deadly and causes a high number of casualties. Model in this project classifies into four different categories using the help of EfficientNet as well as UNet.

The research makes the following contributions:
• A method for classifying four types of cloud forms using satellite photos.
• A smart augmentation strategy based on Albumentation was used to extend the dataset for correct model learning
• Use of a segmentation framework for the task of categorization job to demonstrate that a competent encoder produces productive results when combined with UNet as a decoder.
• Experiment specifics and comparisons have been supplied using the simple usage of EfficientNet models (B0-B5)

The data is from Kaggle's known as "Understanding Clouds from Satellite Images." The photographs include clouds from four different classes: Fish, Flower, Sugar, and Gravel. Due to the tiny footprints, a picture may well be linked from dual orbits of Moderate Resolution Imaging Spectroradiometer (MODIS) onboard these satellites. The dataset was divided 80:20 between train and validation. The training data was equally balanced with 22183 photos, which consisted of 55464 images, with 5546 images in each class. All photos have the same size (14002100 pixels). The Albumentation library was utilized for image augmentation.

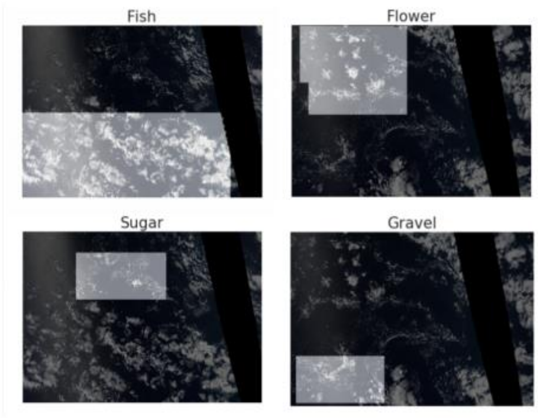


Fig.1 Cloud Patterns

2. Methodology

There were two sets of scores that competed against one another since the dataset was collected from a public Kaggle competition. Scores for the public LB1 are those that are displayed while the competition is still going on.

a. Dice Coefficient:

The Dice coefficient is the evaluation measure employed in this study. The following equation was applied to compare the pixel-wise consistency between a forecasted classification and the corresponding ground truth:

$$Dice = \frac{2 \times TP}{(TP + FP) + (TP + FN)}$$

DSC differs slightly from the most often used assessment metric: model accuracy. They are used to assess the effectiveness of picture segmentation algorithms. The photos are labeled with certain ground truth areas before allowing an automated system to perform the task. The algorithm is confirmed by computing the DSC score, which is an assessment of how analogous the items are derived by dividing the crossover of the two segmentations by the total size of the two objects.

b. Optimizer:

Rectified Adam, or RAdam, is a stochastic optimizer version that incorporates a term to correct the deviation of the effective learning rate. It tackles the bad convergence problem which were previously faced by Adam with fewer epochs. Rectified Adam optimizer was developed by Liu. Et al 2019 [].

c. Loss Function :

As we are having many class or multi class classification, Categorical Cross-entropy (CCE) has been used as a loss function. Categorical Cross-entropy, commonly known as Softmax Loss, is used to compare the differences between two probability distributions.

$$\text{Loss} = -\sum_{i=1}^{\substack{\text{output}\\ \text{size}}} y_i \cdot \log \hat{y}_i$$

Here, yi designates the i-th scalar value in the model's value, where yi is the comparable goal value, and output dimensions specify the quantity of scalar values in the model's output. The softmax activation function is suggested with Categorical Cross Entropy because it modifies the simulation results and tests that it contains optimum factors, as positive responses are desired so that logarithm of every output value yi exists.

Classification loss was calculated through Categorical Cross Entropy and Segmentation loss is calculated using (Categorical loss x 0.7 + DSC x 0.3) respectively.

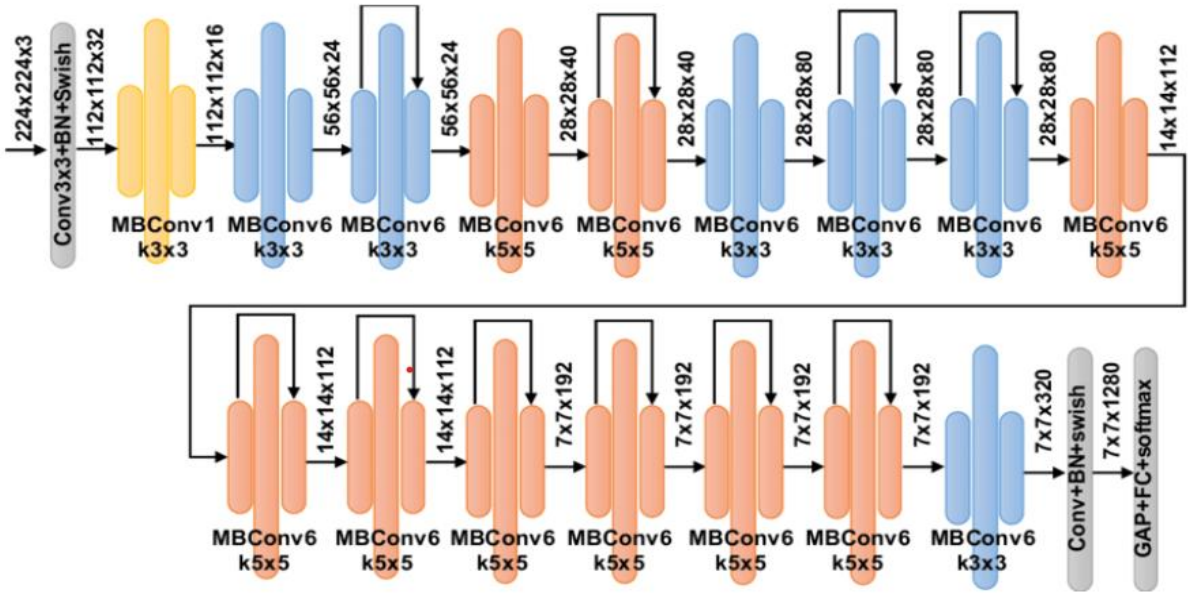


Fig 2. EfficientNet B0 General Architecture []

d. Precision – Recall :

Precision is defined as the ratio of True Positives to all Positives. While recall illustrates the aggregate consistent result, precision represents the fraction of the relevant result. The PR curve is an important assessment statistic since it conveys an additional detailed image of the code's outcome. The X denotes recall, whereas the Y axis indicates accuracy.

When the Area Under Curve(AUC) is high, it means that both are higher. High recall and precision are demonstrated by a uniform false-negative rate and a fewer false-positive rate. Considering cloud formations are complicated, it was critical to determine if the applied model was accurately recognizing the pattern and judging TP, and TN scores, and so the PR curve was a critical assessment benchmark to acknowledge the model's LR.

e. Architecture :

EfficientNet :

It is thought of as a collection of CNN models, although EfficientNet, introduced by Google AI research, performs better than its forerunners [1]. It has eight varieties, numbered B0 to B7, where each model number after that denotes a variation with additional parameters and greater precision.

The three ways that EfficientNet operates:

- Depth-wise Convolution:

Each input channel is handled individually by depthwise convolution. This is a convolution of space. The output of the depthwise convolution on the channel is projected onto a new channel space using pointwise convolution.[1]

- Inverse ResNet:

A layer that squeezes the channel and elongates the channel as well. By this it skip connection to reach channels.[1]

- Linear Bottleneck:

ReLU is included into the final layer in order to prevent information loss.[1]

The other models were not used because they delivered undervalued findings, had poor performance, and consumed valuable runtime as complexity increased. The layers in each of the models (B0 - B7) can be created by using 5 standard modules shown in the figure below.

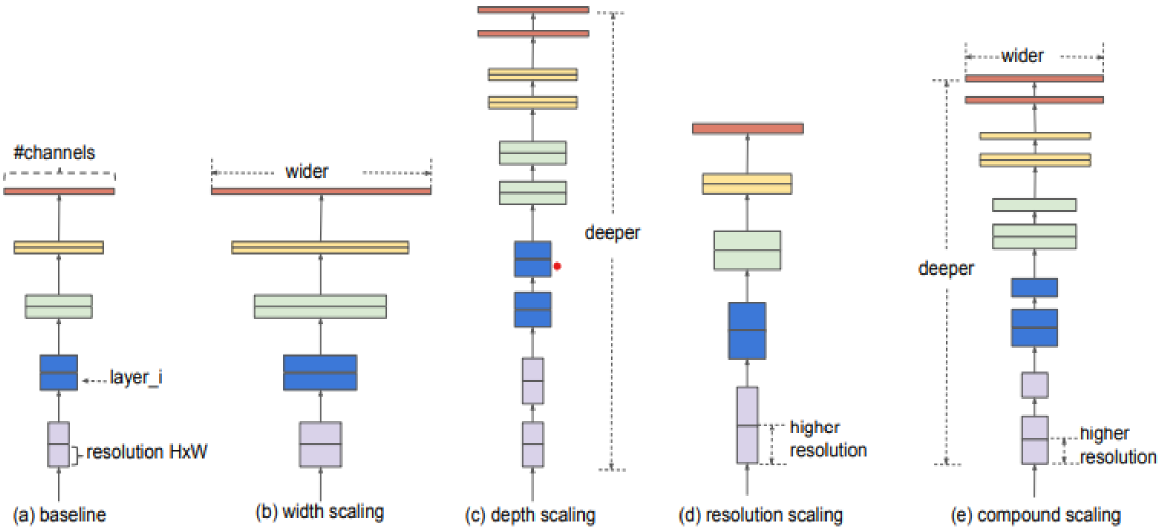
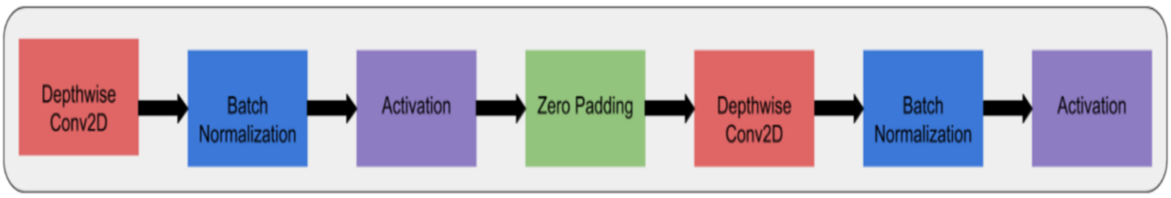
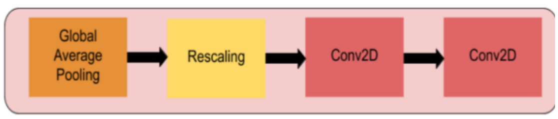


Fig 3 . Model Scaling [10]

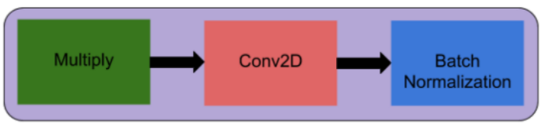
Module 1



Module 2



Module 3



Module 4

Module 1 : Intializations of the first network
Module 2 : Functions as Intializing block for next seven except 1st block.
Module 3 : Skip Connection Block
Module 4 : merges the passed network from the initial network.

Module 5 : creates a skip connection by combining each network section that is connected to its preceding section [1]

The differences between the models are easily discernible, with a progressive rise in the number of sub-blocks[1]. The Inverted Bottle-neck layer, which is an inversion residual block initially used in MobileNetV2 [10], serves as the foundation for EfficientNet. EfficientNet is a dimensioned neural network design in which Compound Scaling, a recently introduced technique, is used to multiply all parameters with a compound coefficient. The definition of scaling up in this context is the methodical, responsible scaling of three components Width, Depth and Resolution.[1]

Every CNN framework is comparable to previous versions. The main difference is that the quantity of parameters is increased by using different feature maps. Compound scaling approach was adopted to scale up in two stages, starting with EfficientNet-B0:

Stage 1: The networks deepest part, breadth, and rs must be searched on a tiny grid since the parameter was set to 1 with the assumption that there would be double as many resources available. [10]

Stage 2: The constants are then kept and the foundation network is scaled up with varying coefficients to produce the subsequent variants from B1 to B7. [10]

Due to the restrictions of the Kaggle notebook, EfficientNet was given precedence in this work.

U-Net

The fully convolutional network is the foundation of UNet.

In order to operate with less training data and provide more precise segmentation, UNet's structure was modified. UNet's purpose is to improve a procuring layer to subsequent layers. Rather than employing pooling operations to replace those layers, upsampling operators are utilized, which yields in output with improved resolution. Based on the inputs, a subsequent convolution layer learns to create the correct output.

This Networks U-shaped design is created via a shrinking and expanding route. A typical convolutional neural network produces a contracting route with each network being followed by a max-pooling operation and a rectified linear unit (ReLU). All through this, feature information has a great rise.

EfficientUNet :

As it can be observed, UNet's contracting and extending route is symmetric. In this, UNet works as a perfect Decoder while EfficientNet works as an Encoder. The Architecture of the EffiencientUNet can be seen in below Fig. 5. The feature map of the encoder's final logit would be first bi-linearly obsolete by a couple of factors, which is then merged with the feature map of the encoder while preserving the same spatial resolution.[11] This operation is continued until the segmentation map reconfigures to the scale of the source images. This framework is asymmetrical, as opposed to the real UNet, and the shrinking path is longer compared to the enlarging path. The use of a resilient architecture, such as EfficientNet, as an encoder, boosted the algorithm's overall effectiveness.

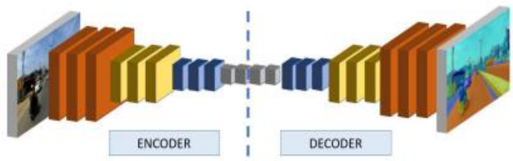


Fig.4 Encoder and decoder Semantic Segmentation [11]

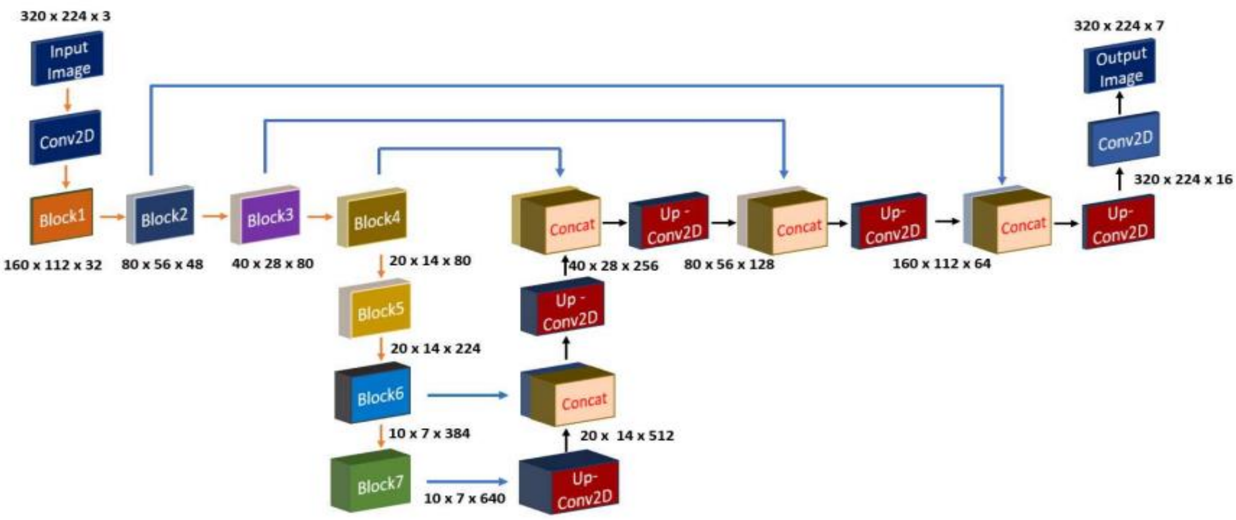


Fig.5 EfficientUNet Architecture with B0 EfficientNet [11]

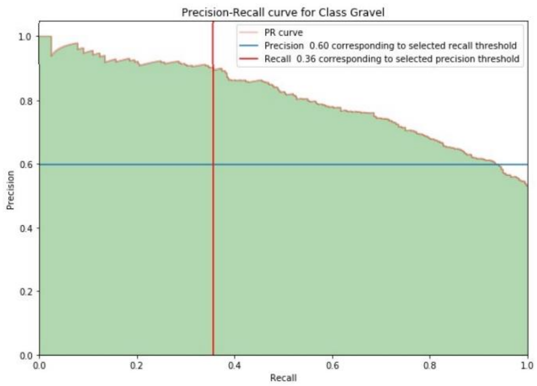
3. Experiment and Analysis :

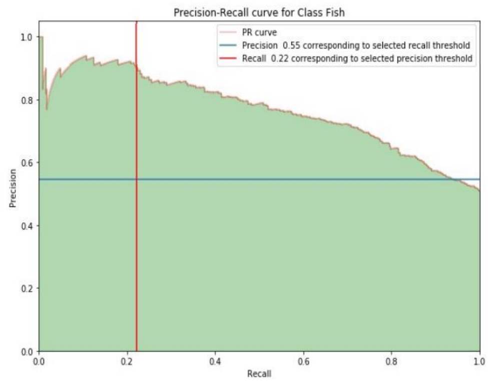
Application of the six various EfficentNet architecture versions that were discussed before has been done as experimentation. For both the initial and fine-tuned versions of all EfficentNet topologies, Mean Validation Accuracy (mVA) has been assessed.

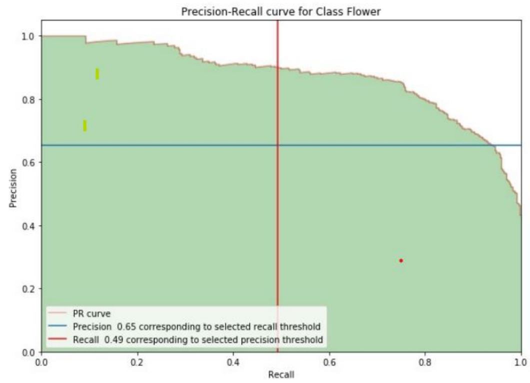
The PR curves are mentioned above. From the above PR curves we can see that Sugar is having the max precision while class fish was having the minimum.

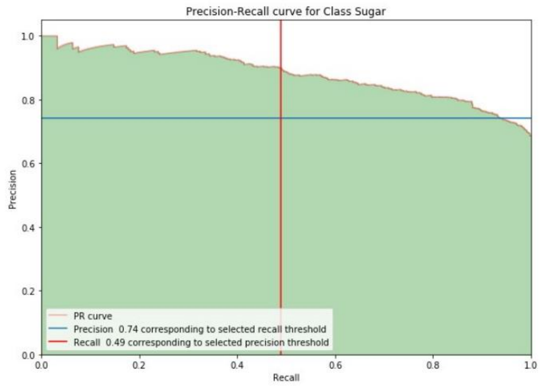
The main aim behind this project was to show that a best classification layer does not always provides the required or desired output, here working the best classification and Merging them as encoder and decoder helped us improve the accuracy of the dataset. Use Of encoder and decoder helped in acquiring better results. Using a Segmentation and classification together , we amplified the accuracy for both.

It is noticed that allowing them to use as Encoder and Decoder, we acquired greater productivity for datasets that need a lot of RAM.









Code :

https://github.com/Priyank2-hub/Assignment-2-DL

4. Conclusion:

In this project, cloud patterns from satellite photos are divided into four categories: UNet was employed as decoder, while six various forms of EfficientNet, from B0 through B5, being utilised as the encoder. The DSC was employed as the assessment metric.

Though EfficientNet was employed in this report, it may have been exchanged with another model that was not evaluated. . The performance of the categorization is also significantly increased by effective picture segmentation. Precsion-Recall Curves were created to illustrate the relationship between accuracy and recall since they indicate the trade-off between the two for different thresholds. In upcoming years , an estimate of categorization and change of the validation set might be incorporated. Since the standard parameters in EfficientNet's complicated designs produced less favourable results, changing these hyperparameters in accordance with the information should enhance the final product.

5. References:

1. Ahmed, Tashin & Sabab, Noor. (2020). Classification and understanding of cloud structures via satellite images with EfficientUNet. 10.1002/essoar.10507423.1.

2. https://rammb.cira.colostate.edu/wmovl/vrl/Texts/SATELLITE_METEOROLOGY/CHAPTER-3.PDF

3. D. D. Turner, A. Vogelmann, R. T. Austin, J. C. Barnard, K. CadyPereira, J. C. Chiu, S. A. Clough, C. Flynn, M. M. Khaiyer, J. Liljegren, et al., “Thin liquid water clouds: Their importance and our challenge,” Bulletin of the American Meteorological Society, vol. 88, no. 2, pp. 177–190, 2007

4. Q. Zhang, J. Quan, X. Tie, M. Huang, and X. Ma, “Impact of aerosol particles on cloud formation: Aircraft measurements in china,” Atmospheric Environment, vol. 45, no. 3, pp. 665–672, 2011.

5. O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in International Conference on Medical image computing and computer-assisted intervention, pp. 234– 241, Springer, 2015.

6. EfficientUNet: Modified encoder-decoder architecture for the lung segmentation in chest x-ray images.
7. https://github.com/zhoudaxia233/EfficientUnet
8. https://www.kaggle.com/code/andersfrgemand/satellite-clouds-u-net-with-resnet-encoder-a-ver
9. https://www.kaggle.com/code/meaninglesslives/nested-unet-with-efficientnet-encoder?scriptVersionId=19966862#Defining-UEfficientNet-Model
10. https://towardsdatascience.com/complete-architectural-details-of-all-efficientnet-models-5fd5b736142
11. B. Baheti, S. Innani, S. Gajre and S. Talbar, "Eff-UNet: A Novel Architecture for Semantic Segmentation in Unstructured Environment," *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 2020, pp. 1473-1481, doi: 10.1109/CVPRW50498.2020.00187.