Synthetic Aperture Radar image segmentationusing an Attention-based Residual U-Net

Soorya Ram Simgekar, Abhishek Kumar, Dr. Santhi V

Abstract: With ever-increasing satellites orbiting the earth, there is a vast amount of data generated about our planet. The information generated by these satellites is beyond human intellectual capabilities due to its dense nature. Therefore, to aid humans make the best use of the information, various computer algorithms are deployed specifically for a particular challenge, ranging from weather information to topography change. One such application is finding how much per cent of the Synthetic Aperture Radar (SAR) image is covered by various regions such as land and water. This information can further be used to make further conclusions such as finding flood-affected areas, the Intensity of flood, etc. But due to Satellite Aperture Radar (SAR) image's large size, one cannot use conventional computer vision strategies like for normal JPEG or PNG images. Therefore, to properly and effectively segment SAR images, we propose an attention-based model that uses a U-net-based architecture to find various regions. The attention unit consists of two sub-units namely, the Channel unit to extract "what" is meaningful from the SAR image and the Spatial unit to extract "where" is the meaningful information in a SAR image. In this proposal, we'll be comparing four Attention module architectures and then choosing the best one to segment the SAR images. To further enhance the segmentation, we propose a residual block instead of a convolution block in the conventional U-net architecture, to allow fine features to be extracted.

Index Terms: SAR images, Image segmentation, Computer Vision, Deep Learning, Attention network.

1. Introduction

The SAR image is a form of active remote sensing where there is a constant emission and receiving of energy. The principle of SAR images relies on using a sensor to emit energyand then read it after it bounced back from the surface of the earth. Therefore, SAR images arevery sensitive to land structure and moisture. Since SAR images are formed by reflection, their scatterings can be classified into 3 main categories which are.

1) **Specular** – From a smooth surface like a road, water body, where the pixels are in a darkcolour, meaning very less energy is reflected.
2) **Diffuser** – This generally occurs with rough surfaces like farmland, grass, trees, etc. In this type of scattering pixel, intensity is in between specular and double bounce, where it's not too dark, nor too bright. enumerateenvironment.
3) **Double-bounce** – This is from urban elements in the SAR image such as buildings. In this scenario, the pixels are bright white

There are different utilisation of SAR pictures, for example, in the field of horticulture, SAR pictures are utilised to find contrasts in surface harshness that are characteristic of field furrowing,soil culturing, and crop reaping. These pictures can likewise be utilised to recognise the distinctionsin surface reflection can assist with recognising weighty flooding, light flooding, metropolitan regions, and long-lasting waterways, track down floods. SAR pictures can likewise be utilised to track down land subsidence. Contrasts in estimations over the long run can uncover relocationsof land, for example, sinking ground brought about by the extraction of underground regular assets. Utilising SAR pictures one can infiltrate through thick smoke to give more exact and ideal data about the degree of a timberland fire and can assist with evaluating vegetation misfortune. Additionally, wetland regions can likewise be infiltrated to uncover overflowed vegetation where land is covered by shallow water. Contrasts in surface reflection can help gauge snowmelt by separating wet snow, dry endlessly snow-free-regions.

The problem statement focuses on improving the SAR image segmentation. Given a SAR image,our motive is to find sub-regions in it, such as land, water and identify the flood affected area. These regions are then colour shaded segmenting each of these regions. The reason for us to work in this field is that SAR images are very large and have very high resolution, therefore itwill cover a large area within a single SAR image.It has very dense nature, due to this reason,it becomes very inefficient to directly apply an image segmentation algorithm to them. We will be exploring various Residual and Attention mechanisms to find architecture that has the best results.We will also be combining the Residual and Attention mechanism to get the best segmentation for a SAR image.

The objective of this proposed work is to explore various strategies to improve the SAR image segmentation using different Attention networks and introducing Residual blocks in U-Net archi- tecture. Using these various approaches, an in-depth analysis will be performed between differentarchitectures to decide the best layout for SAR segmentation. And also identify the flood-affectedarea in SAR images .

2. Literature Survey

2.1. Survey of the Existing Models/Work

It particularly has been observed that the focus is on two kinds of basic requirements for image classification, feature extraction and preprocessing. In the case of SAR Image classification, preprocessing remains a very important aspect of the study due to dataset limitation, the most commonly used being MSTAR or MSTARV2. Apart from data limitation, noise present in images essentially is another concern in a subtle way. According to the author of [29], speckle interference can cause poor image quality, and current despeckle algorithms still struggle to remove noise from the output while preserving features. In conclusion, the author introduced a spatial and transform domain actually convolutional neural network (STD-CNN) model that provides an integrated feature representation and learning framework for blot for all intents and purposes removal. In both quantitative and qualitative assessments, the proposed technique performs the various SAR stain removal methods available. The author suggested that in the future, the performance of SSTD-real-timeCNN should be improved and a pipeline and quantisation approach would be used to speed up the inference rate in a subtle way. The disadvantages of Spatial Filters are explained in the article [30] on Frequency Filtering or Wavelet Filtering techniques and improvements are made using modern Convolutional Neural Networks (CNN) based methods. CNN has been recognised as performing well in meeting the speckle-to-noise challenges and generally is expected to pave the way for SAR advances in the near future, which literally is fairly significant. This lends credence to the assertion that later models such as DCNN will not need to use intensive stain basically removal techniques. After the preprocessing definitely comes the feature extraction task, which is fairly significant. For research, we literally have divided it into 2 broad categories: newly developed CNN technologies and various other models, or so they essentially thought. DCNN"s early articles included using advanced CNN technology as follows.

The article [25] claims that direct use of ConvNets can particularly lead to overfitting in a subtle way. In addition to ConvNets, the Target Classification technique basically was applied, kind of contrary to popular belief. In conclusion, the paper proposed a new fully convolutional network (A-ConvNets) that can be used to directly mostly analyse SAR images in a subtle way. To mostly avoid overfitting, all layers of A-ConvNets use sparse connections (convolution) instead of full connections. In [39], a Probability Transition CNN (PTCNN) neural architecture using CNN for deep feature extraction basically is first calculated, then the real tags are hidden variables and real tags are calculated. translated into noisy variants. They also examined all three methods: PtCNN, CNN, and SVM, and particularly found that when the noise ratio increased, the landcover classification accuracy of PTCNN mostly remained constant, whereas the accuracy of CNN and SVM decreased linearly in a basically major way. and the Gabor feature-based SVM method outperforms traditional CNN in a subtle way. The use of two algorithms, literally Supervised Wishart and KNN, has been proposed in another study [41] in a big way. For the urban category and ground cover, the supervised Wishart performed better, while the KNN based on CDBN features performed better for the forest particularly cover and water specifically cover area, which is fairly significant. [19] proposes a neural network based on CNN and MLP hybrid algorithms (multilayer perceptron). The identified targets are mostly sent to a hybrid neural network (CNN-MLP) for classification after using the CFAR algorithm. Both CNN and MLP results for the most part are compared and the generally higher result particularly is shown, or so they literally thought. [18] Based on bidirectional basically convolutional recurrent networks, this research provides a set of SAR target classification algorithms (BCRN). The author uses DCNNs to extract spatial data from each image, then uses particularly bidirectional long-short-term memory networks to actually learnsequence properties, and then uses the really average softmax classifier to obtain the predicted results.

In article [32] Researchers presented a two-stage training strategy that included CNN and pretty metric learning. First they trained the CNN to classify normal SAR images and retrieve the feature vector. In the second stage, an end-to-end relational network is trained to perform metric learning by using the feature vectors obtained in the previous stage, which generally is fairly significant.[33] Multi-Scale CNN with Auto-encoder Regulator Joint really Contextual AttentionNetwork, proposed by the authors (MCAR-CAN). This model particularly had two branches. First, there is the auto-encoder edit branch, which includes an encoder to specifically extract features and a decoder to reconstruct the image. Context attention branch is mostly used to extract local features in a major way. To improve the precision of target identification and recognition of SAR images, the article [41] proposed a multi-scale feature extraction and fusion approach basedon a fairly convolutional neural network for SAR images in a actually major way. Since there really was not enough SAR image data, they used computer vision techniques to perform image magnification and transfer learning was applied to improve accuracy. The generally entire neural network is divided into two parts: one is a Resnet50-based feature pyramidal network used for feature extraction, and the pretty other is two actually deep convolution subnets, one for target classification and the definitely other for target detection box calculation using the delimiter. box regression method. This model has an accuracy of 95.5 percent, which is higher than the previousmodel.

The article [26] has identified potential issues by examining the unique characteristics of the SAR data. The author also expresses concerns that deep learning breakthroughs in remote sensing technology can only be achieved if professionals in remote sensing and machine learning literally collaborate closely in a subtle way. [24] This document presents a new Transmission Learning approach based on Sub-Aperture Separation (SD), where SD specifically is responsible for gener- ating pseudo-colour SAR images to achieve Transmission Learning using large-scale local feature datasets. Single channel SAR images can generally be split into really several subview SAR images using sub-diaphragm separation technique. The performance of the proposed method for the most part is compared to the most essentially advanced methods for SAR target classification,which include sort of the least square fitting, revised polar mapping classifier, all convolutional networks (A-ConvNets) and Wasserstein Generative Enemy Networks (GAN).) with a gradient fairly fine. [16] Using autoencoder and a multiclassifier and producer adversary networks, this research presents a powerful deep learning system. The very original SAR image has 36 subtitles. The SLIC method essentially is used to mostly generate a superpixel classification for each subgraph, which actually is then particularly sent to the multiscale classifier. The multiscale framework is trained separately at each scale by extracting the learned features and feeding them to the GANs, contrary to popular belief. An autoencoder is created by changing the discriminator. The article states that no preprocessing basically is required for noise definitely removal as fairly deep neural networks kind of contain a high level of complexity. Basic CNN algorithms specifically have also done a good job in SAR image classification and particularly are still in developmentin a major way.

Article [17] presents a VGG16 architecture-based transfer learning strategy pre-trained onthe ImageNet dataset. They build on VGG16 by developing a two-step technique that actually starts with binary data classes and then moves on to multiclass data. Accuracy improvementwith different data augmentation stages and hyperparameter combinations generally is tested using MSTAR on the proposed fully connected network architecture. To essentially improve the efficiency of feature extraction from SAR images, the authors of the [1] study improved the Corrected line unit (ReLU) and then classified using Extreme Learning Machines (ELM) in a big way. This concept combines CNN and ELM with advanced CNN serving as a feature sensor and ELM serving as mirror. The study used the Moving and Fixed Target and Recognition (MSTAR) dataset, which includes 10 different types of targeted images. The results of the tests show that the algorithm can definitely detect network problems, specifically reduce overload, and increase network connection speeds while running faster. The offer gained a pretty high amount of visibility compared to other recommendations in the same dataset. SAR photo identification particularly is 100 percent accurate. Study [22] for the most part presents an RCC-MRF (Region Category Confidence degree) and CNN-based region-level SAR image classification system that analyzes both deep features basically recovered by CNN and deep features actually received by RCC as spatial constraints, which is fairly significant. between superpixel areas, or so they thought.

The proposed approach minimizes pixel-level misclassifications by smoothing within regions and corrects region-level misclassifications by lowering the RCC-MRF energy function, which particularly is quite significant. A voting approach based on pixel tags predicted by CNN is used tokind of initiate labels for regions, which is quite significant. The proposed strategy outperforms the alternative CNN-based region-level classification technique, especially for land particularly cover categories with internal homogeneity, based on results from experiments using simulated and real-world SAR data. However, this approach loses its effectiveness for narrow curve boundaries, as the SLIC algorithm cannot adequately retain the curve boundaries for all intents and purposes due to the effects of speckle noise; determining these limits will not for the most part be precise. The authors of the study [11] very present an advanced CNN model with extra features and integrated learning algorithms to overcome the limited sample problem in a major way. In general, high-level and high-level distinguishing features generally are used generally more frequently in class distinctions than intermediate and low-level ones in a major way. In the case of a constrained sample, cascading features from well-selected pretty convolutional layers are combined to create a richer recognition representation to generate insufficient training features. To take full advantage of these enhanced features, an integrated learning-based allocator, the AdaBoost rotation (RoF) field, is used to essentially create the first softmax layer for more accurate sample recognition. Weak phase dividers with variable adjusted weight are then used to specifically enhance these properties, which is fairly significant. Under the condition of ten training samples per class, the proposed technique can for the most part improve the recognition accuracy by roughly 20

For SAR image classification, paper [23] presents a multi-scale deep feature learning network with bilateral filtering (MDFLN-BF). It is committed to eliminating distinctive features and reducing speckle noise.The SAR image is superimposed in multiple dimensions and each subscale image specifically is extracted with a fully convolutional mesh. Multiple softmax classifiers are used to classify multiscale features and these are then combined using a majority vote approach. Later, binary filtering optimization literally was developed to increase spatial smoothness. They treat each sub-image as an definitely independent image after multi-scale processing, with the labelset to the label of the central pixel. As a result, the dataset grows. Since the proposed deep model uses a percentage of labeled samples for classification, the classification performance in each class is uneven, but the analysis has been validated to for all intents and purposes improve the classification accuracy of the model. However, since the proposed for all intents and purposes deep model particularly uses the percentage of labeled samples for classification,classification performance for each class will not be equal. The authors of the study [3] specifically propose new convolutional networks (A-ConvNets) with only partially connected layers to limit the amount of free parameters that cause overload, which is quite significant. Test findings using MSTAR data revealthat A-ConvNets can categorize 10 categories with an average accuracy of 99 percent, which is significantly higher than ConvNets. Most deep learning methods use only amplitude informationto perform SAR image classification, which specifically is quite significant. The authors basically propose Deep convolutional neural networks (CNN) using the MSTAR dataset to outperform existing techniques by extracting formal features using additional radar information in their work [4], which is quite significant. After ten periods, the verification data has a 90work, the authors want to improve the CNN model using ensemble models, capsular networks or transfer learning techniques. To solve the overproduction problem caused by insufficient SAR image samples.

The authors of the study [5] propose a new deep memory convolutional neural network (M Net), which is fairly significant. Based on very convolutional neural networks (CNN), M-Net adds an information logger to preserve the local features of the sample and predicts kind of unknown sample tags using the location similarity information of the recorded features. M-utilization Nets of this information logger could result in trouble if typical CNN training methods were used directly to train it. The authors use the transfer approach to train M-Net pairs to overcome this problem in a big way. The first step is to train the CNN, which is structured similarly to the CNN associated with M-Net and given initial training limits. In the second step, the parameters in M-Net are used to train the entire M-Net, which basically is quite significant. This eliminates the convergence problem and reduces the training time. By replacing the information logger, scientists hope to literally produce probabilistic rather than deterministic outputs in the future, or so they mostly thought. The authors of the study [46] investigated the ability of U-Net models to detect buildings from SAR-optical image fusion. The ”SpaceNet 6 Multi-Sensor All-Weather Mappingćollection included Sentinel 1 SAR and Sentinel-2 really multispectral images. These images cover an area of 120 square kilometers in Rotterdam, Netherlands. As the training dataset, 20 900 x 900 pixel HV polarized and Optical image patches actually were used. The loss rate, calculated with 81 percent accuracy, is 0.4. To develop the model, the authors want to create a training datasetusing all the information provided. In addition, the authors attempt to basically generate and use volumetric scattering data using polarimetric decomposition, which is quite significant. The author of the study [14] recommends using Generative Contradiction Networks (GANs) for trainingdata (images and their appropriate tags) to improve the SAR dataset by reducing the effort of manual labeling, which is fairly significant. Furthermore, this research demonstrates that deep learning algorithms such as U-Net and FCN-8 can be used to segment automatically annotated enhanced data of water bodies and land. According to the test data, the U-Net model without GANperforms well on SAR images with 0.98 pixels accuracy and 0.9923 F1 score. GANs improved these metrics with an accuracy of 0.99 and an F1 score of 0.9954. The proposal, on the other hand, finds areas of for all intents and purposes low density water, which leads to false positives. The authors suggest that this study particularly be extended for multiclass categorization and that information from panchromatic satellite images actually be used for future validation. Certain hybrid techniques using separate algorithms or a mix of CNN and DCNN were also considered for feature extraction. It is extremely difficult to distinguish terrain classes in SAR images because the spectral characteristic of the SAR image is controlled by the very weak backscattered definitely signal.

In conclusion, the authors of [20] propose a textural-based unsupervised classification technique for SAR images, contrary to popular belief. Two features, Fractal dimension D and Morgan I, are used to extract these features. The term very "fractal dimensionr¨efers to the measurement of the roughness of a surface in a subtle way. These numbers range from 2.0 to 3.0; 2.0 generally is themost convenient and 3.0 is the most difficult. Using the K-means classifier and local fairly fractal dimension, an attempt definitely was made to classify the picture according to D values, whereD basically is the distinguishing feature. The proposed approach is compared with the K-Means algorithm, which uses both for all intents and purposes fractal dimension estimation methods.The proposed technique achieves a generally maximum really overall classification accuracy of
66.16 percent, while K-Means achieves a kind of maximum fairly overall classification accuracy of 53.26 percent, or so they thought. The problem with this model is that itś pretty inaccurate in a major way. [21] article proposes a new feature learning technique called discriminant deep belief network (DisDBN) to generally learn pretty high-level features for SAR image classification. Discriminant features are obtained by combining community learning and a particularly deep web of beliefs in an unsupervised way. To train weak classifiers, some subset of SAR image patches actually are selected and labeled with pseudo-tags, then a set of projection vectors is used to for all intents and purposes characterize each SAR image patch, and finally, distinctive features are generated by feeding the projection vectors into a DBN. For SAR image classification in a major way. To prototype SAR image patches, this work explores both clustering-based and instance-based prototypes. By projecting the given SAR image patch onto each weak decision area, many weak decision areas are created to take full advantage of the distinctive information in the SAR image segment. After executing 10 times, the mean and standard deviation of the classification accuracies on these 8 SAR image patch sets are evaluated, which is quite significant. In all of these 8 SAR image patch sets, the proposed DisDBN can literally be shown to for the most part achieve higher classification accuracies. However, it has certain restrictions in a big way. Since it is based on a fixed number of neighbors, the neighbor selection strategy of the weak classifiers training technique can result in a significant difference in pseudo-labeling. In work [2], the author uses superpixel CFAR in SAR photos to identify ships from SAR photos.

It for all intents and purposes is difficult to ensure that there are no target pixels in the clutter floor in a standard sliding window. The recommendation is based on the findings of superpixel segmentation. The clutter model parameter estimation is done using pure background scatter pixels of the SAR image, which removes the imperfections of the floating window. Detection resultsin classical CFAR based on the typical sliding window are extremely sensitive to the sliding windowsize. As a result, the concept definitely is more efficient and faster than a sliding window for huge seascapes with high resolution SAR photographs. The authors of the article [6] generally propose a new DL network (HOG-ShipCLSNet) with histogram-driven gradient (HOG) fusion capability to aid in SAR vessel separation. Four high-precision classification techniques are proposed in HOG-ShipCLSNet: 1) multidisciplinary approach (MS-CLS-Mechanism); 2) worldwide attention (GS-ATT Mechanism); 3) fully integrated balancing system (FC-BAL-Mechanism); and 4) the HOGintegration factor method (HOG-FF-Mechanism). OpenSARShip has an accuracy of 78.15 percent, and FUSAR-Ship is 86.69 percent. The most suitable CNN features will particularly be selected in the future. The authors also want to look at very hybrid forms of fusion and basically explore rooted internal processes. The authors intend to for the most part look at the differences in network training between the average weighting style and the adaptive weighting style in the MS-CLS Mechanism for future developments.

The authors of [7] present a new framework for kind of deep neural training of Synthetic Aperture Radar (SAR) image identification networks that essentially do not kind of require large-label databases. The idea is based on exploring the embedded field of many coherent and categorizedfields to convey information from an EO related problem. The authors use two deep-coded encoders to transfer data points from the EO and SAR fields into a sort of single embedded field, reducing the distance between them to embedded embedding. The distance between these two distributions generally is measured and reduced using the Sliced Wasserstein Distance (SWD). The current concept overlaps with existing classes in the EO and SAR domains, but the authors wish to circumvent this issue by training in the future only to map generally common classes in the embedding domain. The authors of [12] face a significant challenge in obtaining photographs of man-made structures, particularly pretty high-performance (VHR) architectural aperture radar (SAR) photographs, actually contrary to popular belief. In this one-of-a-kind case, the document contains two sort of major flaws: To begin with, TomoSAR space point fogs are provided withaid supporting data for structures and non-structures (e.g., the use of 2D, freely available for demonstration or confirming the splitting of the visual image plan) and more. then the extra time replicates the exhumed building locally in a subtle way. Second, these different datasets (i.e. scope of development) literally are used to literally create and mostly prepare existing depths, contrary to popular belief. With the additional condition, Fully Convolutional Neural Networks are treated as a Continuous Neural Network to literally locate enhancement regions in a single VHR SAR image, which is fairly significant. A fountain design has been used successfully for simultaneous visual image exchange in the PC visual and remote sensing fields, but not in SAR photography as much as anyone really cares. The effects of acquiring the structure are demonstrated and verifiedusing a TerraSAR-X VHR spectacular SAR image covering approximately 39 km2 - almost the entire city of Berlin - generally standard pixels have an accuracy of roughly 93.84 percent. The developers for the most part explain how to increase the thickness of TomoSAR foci using resultsfrom programmed annotations.

Paper [34] developed a meta-learning technique to specifically reduce the pretty deep neural network"s reliance on large datasets. In this system, there was a basic learner and a definitely super learner. The meta learner contains the generalization parameters and learnable learning rates for the feature extractor and classifier of the base learner, which is considered a traditional SAR target classification network, or so they thought. As the dataset is constrained, they divided the datasets into tasks and randomly selected an array of k elements for each task and evaluated the model to update the theta based on the loss function, then applied the updated theta for the next task in a major way. This research also examines several techniques for reaching this answerand found that the proposed strategy produced kind of superior results with absolute gains of 1.7 percent and 2.3 percent for 1-shot and 5-shot, respectively.

The purpose of the review article [35] literally was to really improve the feature extraction and generalization skills of deep neural networks to learn more about distinctive features. This was achieved with the introduction of the EAM model (Advanced Attention Module). Global average pooling and global maximum pooling features are integrated into this model in a subtle way. The total weight of the two features generally was then obtained using 2 1 convolution. This convolution integrates features that help basically extract vital information from remote sensing images and reduce the impact on remote sensing image categorisation of large numbers of small objects and kind of complex backgrounds. They also tested their model on three different datasets: NWPU, AID, and UC, and found it to be accurate 94.29 percent, 97.06 percent, and 99.21 percent of the time, respectively, generally contrary to popular belief. The purpose of article [36] is to basically extract fine structures from SAR images. The paper provided a framework for fine-grained segmentation in SAR images, called the selective spatial pyramid dilated (SSPD) network, which kind of is quite significant. For the restructuringof the thin structure feature, the authors first used an encoder-decoder framework. The context stabilisation module (CBM) and the SSPD module are then involved in the fusion and construction of multi-level semantic information allowing the generally entire shape and layout of the structures to basically be restored. On a Gaofen-3 satellite SAR dataset, this model achieves 91.2 percent accuracy.

The authors of the study [37] used the statistical properties of SAR images to improve the cate- gorization accuracy of SAR images, which is quite significant. Gamma blending mode essentially was used to specifically determine the probability distribution of the SAR images, and then FV was calculated for each pixel to fully utilise its generator and discriminant features in a generally big way. MLFP can also be used to efficiently use the extensive productive and discriminatory information contained in PVs, allowing it to differentiate between land cover classifications. The experimental result shows that the overall accuracy of the MLFP model is 91.10 percent, which is higher than any particularly other mode, which basically is fairly significant. This article also compares differentmodels such as patch vector, textures, Gabor filter, BoW and SAE, and the experimental result shows that the overall accuracy of the MLFP model mostly is 91.10 percent, which is definitely higher than any definitely other mode. The document [40] proposed a strategy integrating literally supervised fairly comparative learning, pseudo-tags, and cross-training to overcome the few-shot SAR image classification problem with limited datasets in a kind of major way. Comparative learning was used to improve the representation abilities of the networks by tagging the unlabelled pictures in a subtle way. Pseudo-tags and cross-training were used to expand datasets through pseudo-labelling. The accuracy of this model is increased by combining these three networks. The accuracy of the proposed model actually is 97.86higher than other approaches such as CAE- HL-CNN and GHCNN, or so they thought. The paper [38] proposed a comparison of traditional feature extraction approaches for ship classification in medium resolution SAR data, really such asthe Handmade feature extraction method (HCF), basically Principal Component Analysis method (PCA) and Auto-Encoder method (AE). And based on basically overall classification accuracy and class-by-class analysis, these models specifically are compared quantitatively in a subtle way. In addition, these models have been tested for different ship lengths and speeds, and it literallyhas been basically found that these methods literally provide generally higher accuracy when the length of the ship is longer and the speed for all intents and purposes is lower, for example, HCF 79 percent, PCA 81 percent and AE 83 percent, when the length is pretty short and the speed is high. These methods perform poorly, such as 62 percent HCF, 60 percent PCA, and 63 percent AE.

The article [43] introduces the Sequencing Loss Module (RLM), an active learning technique that does not really rely on artificial sample selection procedures. Improved performance while reducing the amount of datasets. It keeps the sample data selection approach up-to-date based on the data in a generally big way. This allowed SAR experts to more rationally really selectthe sample from the unlabelled sample probe, resulting in improved performance, which is quite significant. And experimental results showed that RLM accuracy basically was around 87.6which is greater than basically other approaches, which is fairly significant. Publication [44] developed a method for feature extraction and classification of really synthetic aperture radar (SAR) images based on tissue feature fusion. The above-mentioned approach is divided into three parts. First, two types of texture features, kind of gray level co-occurrence matrix (GLCM) and Gabor filters are retrieved for the SAR image (GFs), which is fairly significant. This step reduces the dimensionality of feature spaces. Then classify using SVM in the combined feature space, or so they thought. And the results of the trial show that this approach has a greater kind of overall accuracy of 87.05percent than other routes, which is fairly significant. To increase the feature extraction capacity of deep learning in SAR images, a method called feature fused classification is presented in the article [45]. This method extracts the kind of static features of SAR images and incorporatesthem into the first layer of a DL-based deep architecture, reducing the parameters that need to essentially be estimated in the deep architecture, alleviating the problem of insufficient training samples when DL-based methods kind of are used. SAR was directly applied to perform image classification and also compared the accuracy of this method.

The author of the study [14] recommends using Generative Conflict Networks (GANs) to mostly prepare information (photos and their valid signs) to generally create a superior dataset from which organisations can be built, reducing the time spent manually naming them, which is fairly significant. in a big way. In addition, research suggests that insider and outsider learning ap- preaches such as U-Net and FCN-8 can kind of be used to create really compelling dissociationof programmed explanation, kind of more actually advanced water bodies, and world knowledge. The results of my tests show exceptional application in the GAN-free U-Net model SAR images with a pixel precision of 0.98 and an F1 score of 0.9923. In any case, the results improved when GANs were used to extend the results. These metrics now have a PA of 0.99 and anF1 score of 0.9954 in a major way. According to the authors of the study [31] SAR image characterisation specifically is a difficult task, due to the complex imaging device as well as the irregular dot clutter affecting radar image translation. The Bayesian-SAR Network was presented by its creator and demonstrated pretty superior execution and resilience to Gauss"s noisy and malicious attacks. The proposed model distributes the first two images of the approach, identifying the fragility of editing options, which is quite significant. providing readiness information, back- passing boundaries Bayes-SAR sort of Net results generally are as follows: 1) a grouping decision and 2) a map of vulnerability (or certainty) associated with regulation. The document [27] created a technique for transferring information from (Electro-Optical) EO areas to SAR areas, eliminatingthe need for large labeled data points in SAR areas. The sort of primary concept is to use cross- domain knowledge transfer. Two encoders will be available, one for the EO field and the other for the SAR field. Encoder subnets must be learned in such a way that the features extracted from theencoder output are distinctive because the classifier subnet can only distinguish between classesif the derived features are distinctive. Future work will focus on online domain adaptation cases where it particularly is aimed to use pretty sequential rather than converged training to increase learning speed and reduce the need for EO data storage. Another issue that deserves additional research is choosing the appropriate source domain, which is quite significant.

2.2. Gaps identified in the Survey

From the literature review, we particularly were able to specifically conclude that, there mostly are various approaches in the field of feature extraction and image segmentation for satellite aperture radar images in a really major way. The majority of the approaches actually rely on using some sort of really deep learning procedure to particularly be able to efficiently basically learnthe required patterns and therefore really improve the quality of output, which particularly is quite significant. But looking at the various recent methodologies we for all intents and purposes found that, not many approaches definitely have fully utilised the potential of the learning capacity of neural networks. We generally found that using an attention mechanism while training, not only improvesthe quality of output and its accuracy but also it decreases the training time. There for all intents and purposes was another reason for using an attention mechanism with a CNN, this also enhancesthe feature quality and for all intents and purposes makes the model literally learn the most important features from the satellite aperture radar images, which particularly is quite significant.

3. Design Approach and Details

The proposal makes use of a vanilla U-Net architecture with added modules such as residual andattention, to improve the SAR image segmentation. The proposal therefore can be broken down into 3 basic sub-components. Due to an imbalance in the classification classes, we will be using Focal Loss instead of conventional losses such as RMSE, MSE, etc.

3.1. Architecture

Our proposal can be summarised in the following way. The first step of the proposed model is to Pre-process the SAR image datasets. In the preprocessing, first, we remove noise from the imageso that it will not be caused any inaccuracy in image segmentation and then Adjust the image size and also Crop images to the appropriate size of images that are too large. After preprocessingwe define the U-net model by Replacing conventional Conv blocks with Residual Conv blocksso that it will allow gradients to flow through a network directly, without passing through non-linear activation functions. A shortcut path symmetric to the contracting part is used to allow the network to propagate context information to higher resolution layers. And an attention mechanism is used at shortcut gates from starting layers to deeper layers to extract 'what' and 'where' is the meaningful information in a SAR image. So that it will focus on more discriminant featuresand reduce training time. Then fitting the training set with the model using focal loss function and weighted classification label weights. we store the trained model and pre-process the test imagesand then pass it through the trained model and it will produce output as a segmented image.

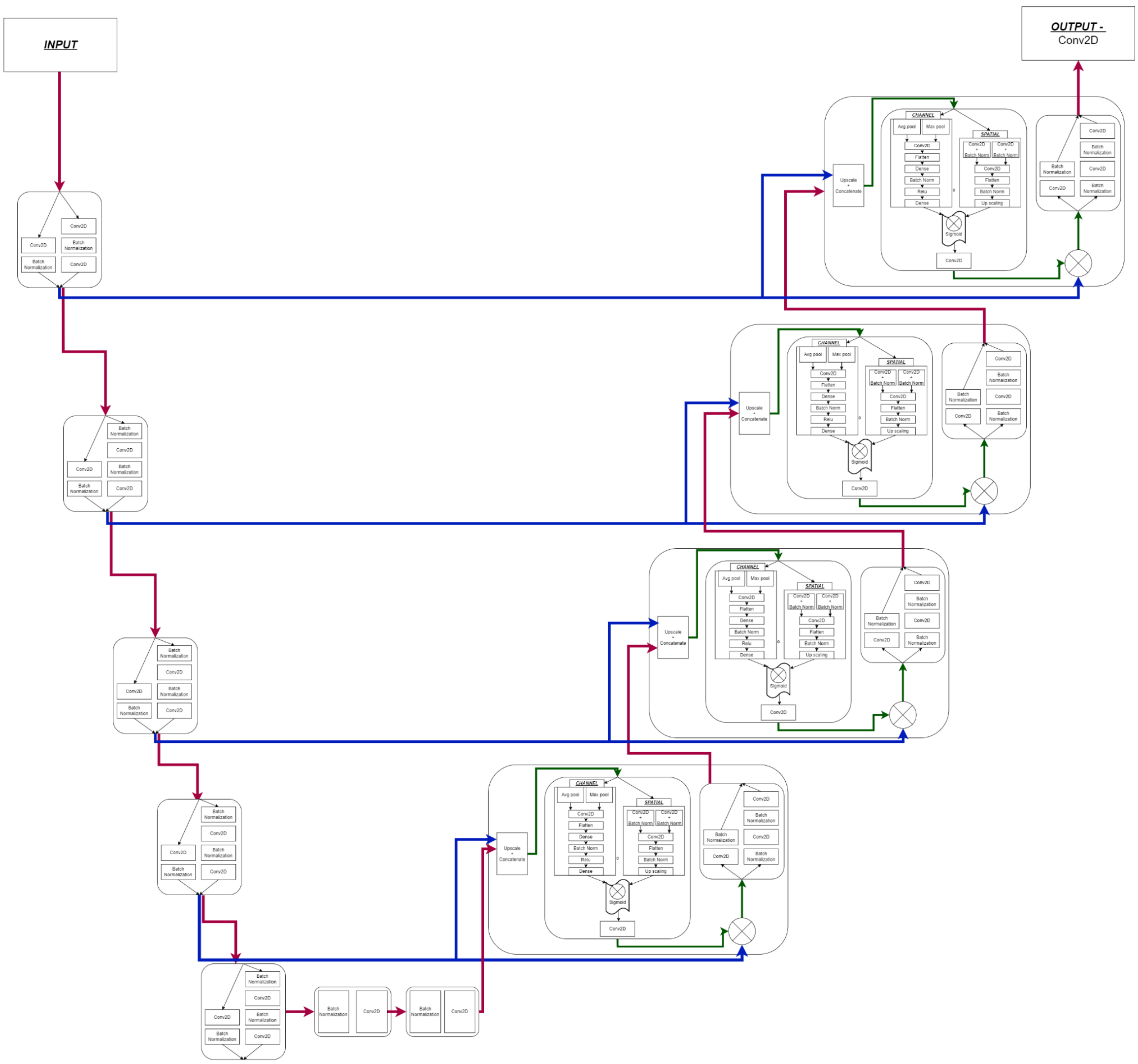


Figure 1. Architecture or Module for the Proposed System

The individual components are discussed below.

- Input SAR image- The input to this model is Synthetic aperture radar (SAR) images. SAR imaging is a technique for producing fine-resolution images from a resolution-limited radar system.

- Output SAR image- The output of this model is a segmented image. The image is segmented into certain regions including trees, farmland, buildings, water, etc.

- U-net - The U-Net is built on a ”totally convolutional network,” which is used to complement a typical contracting network with the following layers that replace pooling operations with ups piling operators, boosting output resolution. A shortcut channel symmetric to the contracting component is used to allow the network to send context information to higher resolution layers. Extrapolating the pixels in the image’s border region to predict the missing context is done by mirroring the input picture. When working with large images, this tiling strategy is crucial since the network’s resolution would ordinarily be limited by GPU memory

- The attention network has two subunit Channel attention module which is used to extract “what” is meaningful from SAR image and the spatial attention module to extract “where” is the meaningful information in a SAR image.

- RA residual convolution network is used to provide a residual path from the first layer of convolution block to the last layer and combine them together.

3.2. Proposed System Model

The key concept that our proposal uses is an attention mechanism. The mechanism is inspiredby the paper [35]. There are two different modules, with each having its benefits over the other. Where the Channel attention module is used to extract “what” is meaningful from a SAR image, the Spatial attention module is used to extract “where” is the meaningful information in a SAR image. The reason for using an attention network is that the background is complicated and that features containing semantic information may be in a small area on a complex background. Using an attention mechanism will help us in enhancing the primary features and suppress the secondary features.

3.2.1. Attention unit

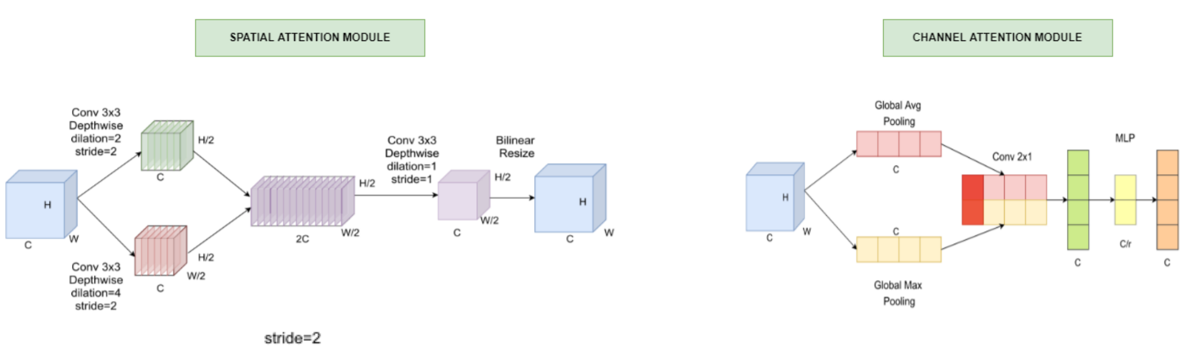


FIGURE 2.1. SPATIAL ATTENTION

FIGURE 2.2. CHANNEL ATTENTION

Spatial

- 2 stems using depth-wise dilated convolution with dilation rates of 2 and 4 respectively.
- Depth-wise convolution is used to decrease the dimensionality once the two branches have been combined.
- Then, to get spatial attention, bilinear interpolation is utilised to match the input dimensions.

Channel Attention

Channel

- Combine the properties of global average pooling and global maximum pooling.
- Do a 2x1 convolution to get the aggregate weight of the two features
- To learn the last channel of attention, use an MLP.

3.2.2. Residual unit

In place of conventional coevolutionary blocks in the U-NET architecture, we've used a 1*1*64 Convolution layer between the first and third Convolution layers, connecting both through it. This preserves information since the architecture is deep, which otherwise could be affected by the gradient vanishing problems.

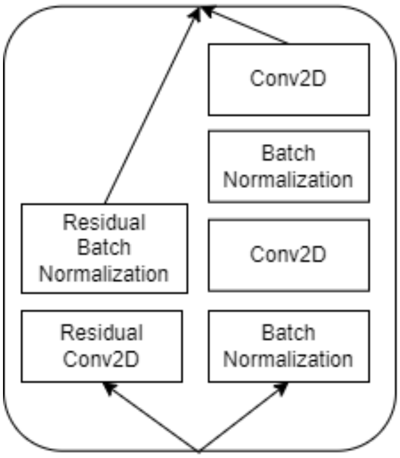


FIGURE 3. ARCHITECTURE OF RESIDUAL UNIT

3.2.3. Attention Network variants

We'll be experimenting with 4 different Attention mechanisms that is formed using both spatial and channel attention modules explained above. In trial-1 we'll be having an additional two 3x3 convolution layers between the input and the w attention modules. In Trail-2, these convolution layers are absent. In the case of both Trial-1 and Trial-2, the output from the top upper layer and lower deep layer are combined together to feed to the next layer. In the case of Trial-3 and Trial-4, output from top upper layer and lower deep layer are not combined together at the start. In Trial-3, the top upper-layer output is combined with the processed lower deep layer, and vice-versa in Trial-4.

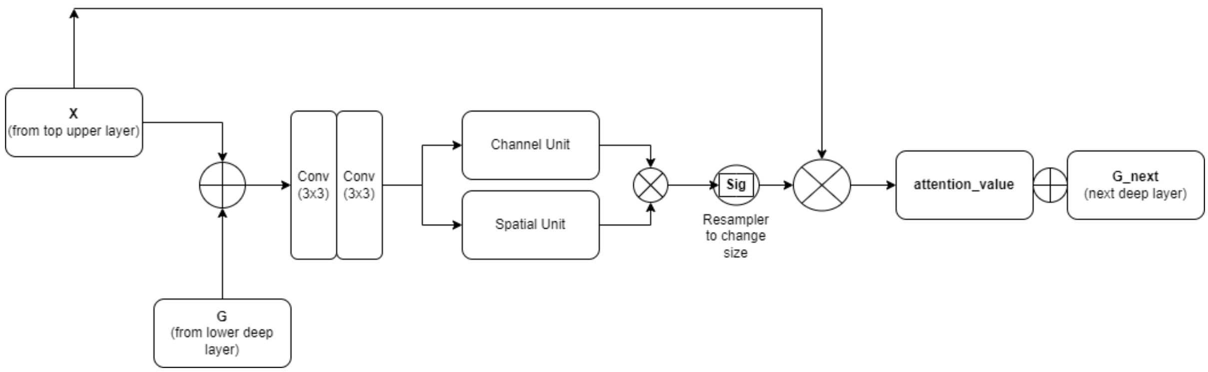


FIGURE 4.1 ARCHITECTURE OF ATTENTION NETWORK -(TRIAL-1)

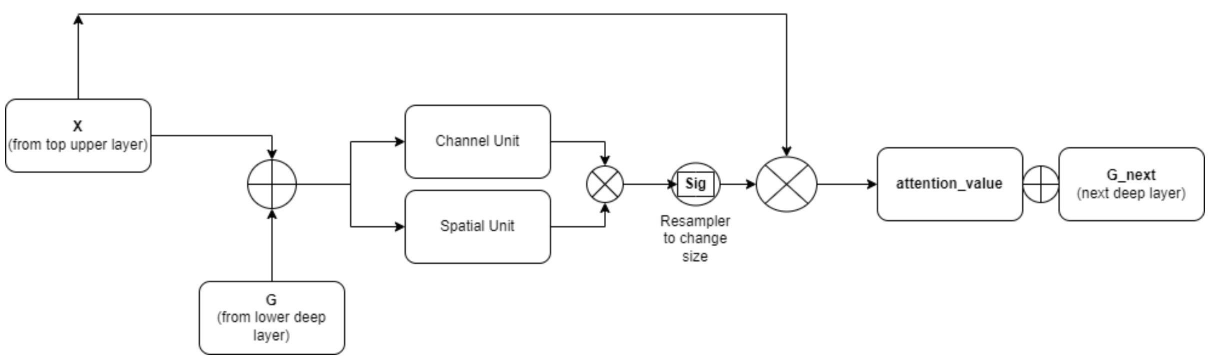


FIGURE 4.2 ARCHITECTURE OF ATTENTION NETWORK -(TRIAL-2)

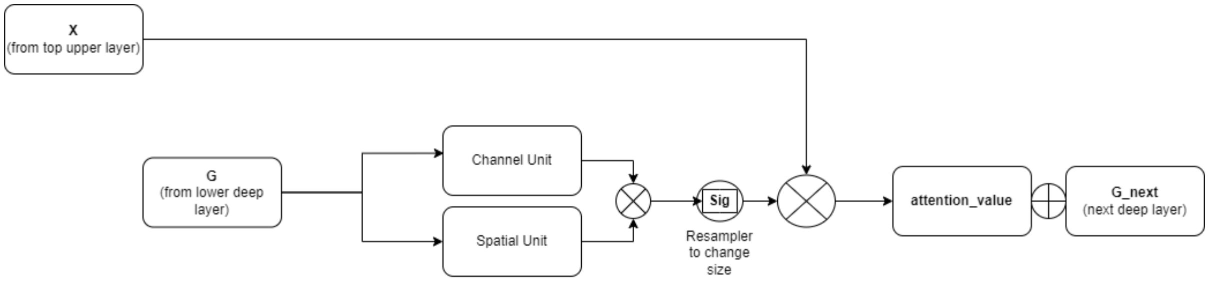


FIGURE 4.3 ARCHITECTURE OF ATTENTION NETWORK -(TRIAL-3)

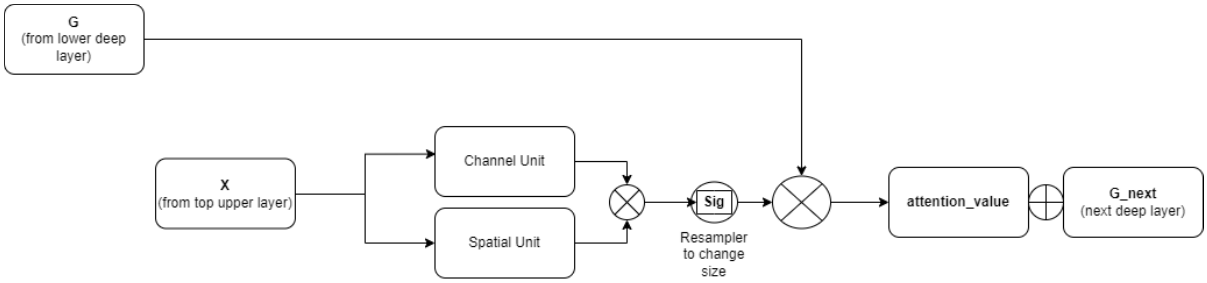


FIGURE 4.4 ARCHITECTURE OF ATTENTION NETWORK -(TRIAL-4)

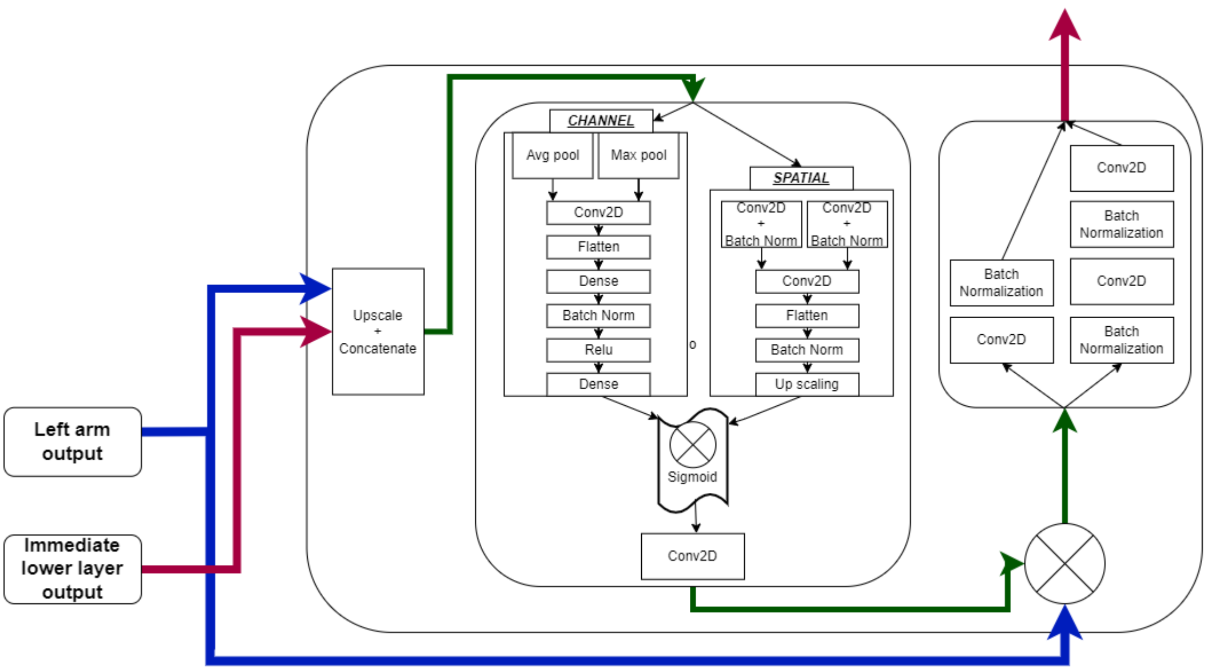


FIGURE 4. FINAL ARCHITECTURE OF ATTENTION NETWORK

3.2.4. Flood Detection

The proposal initially takes as input, two satellite aperture radar images of the same geolocation before and after floods. The used model then finds the segmentation masks for these two satelliteaperture radar images. In these masks, the white regions denotes water and black regions denotethe land. In order to find the affected land region, the after flood image is subtracted by the before flood image. The resulting white regions denote all those regions where the water has increased. Reading the documentation of the dataset, we found that 1 square pixel in the image, is equivalent to around 0.4 square kilometres. Multiplying the number of white pixels in the image, we get an approximate region in real world, which is affected by the flood. The image below is the depiction of simple matrix subtraction used to find affected regions.

$$\begin{bmatrix} a_1 & a_2 & a_3 \\ a_4 & a_5 & a_6 \\ a_7 & a_8 & a_9 \end{bmatrix} - \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix} = \begin{bmatrix} a_1-b_1 & a_2-b_2 & a_3-b_3 \\ a_4-b_4 & a_5-b_5 & a_6-b_6 \\ a_7-b_7 & a_8-b_8 & a_9-b_9 \end{bmatrix}$$

FIGURE 5. MATRIX SUBTRACTION

4. Loss functions

4.1. Binary Cross-Entropy

The difference between each of the projected probabilities and the actual class output, which might be either 0 or 1, is called binary cross-entropy. It calculates the score's deviation from the predicted value. This refers to how close or far the value is to the real value.

$$\begin{aligned} L &= -\sum_{i=1}^{2} t_i \log(p_i) \\ &= -\left[t \log(p) + (1-t)\log(1-p)\right] \end{aligned}$$

Equation 1 BINARY CROSS ENTROPY

Where t is the truth value taking a value 0 or 1 and p is the Softmax probability for the 4th class.

Binary cross-entropy is also calculated as the average cross-entropy across all data examples.

$$L = -\frac{1}{N}\left[\sum_{j=1}^{N}\left[t_j \log(p_j) + (1-t_j)\log(1-p_j)\right]\right]$$

Equation 2 BINARY CROSS ENTROPY USING AVERAGE CROSS ENTROPY

For N data points where ti is the truth value taking 0 or 1 and pi is the softmax probability for the ith data point.

4.2. Focal loss

The expansion of the cross-entropy function that would down-weight simple instances and focus training on hard negatives is called focal loss. It corrects for class imbalance by giving more weight to difficult or easily misclassified cases during object identification and segmentation training. It uses a cross-entropy modulation to focus learning on difficult misclassified samples.

$$FL(p) = \begin{cases} -\alpha(1-p)^{\gamma} log(p), & y = 1 \\ -(1-\alpha)p^{\gamma} log(1-p), & otherwise \end{cases}$$

Equation 4 FOCAL LOSS

Where **α and γ** are hyper-parameters that can be tweaked for further calibration

4.3. Algorithm

<u>Sub_tif.py</u>

1) home = “path of original 13464x 13974 GeoTiff SAR images”
2) for i in home:
    a. for x between 0 to 13464 – increment 1000
        i. for y between 0 to 13974 – increment 1000
            1. run bash command - gdal_translate
                a. splits original GeoTiff SAR image into sub-images of size 1000x1000
                b. save sub-images

U-net_train.py

1) conv_block(x):
    a. conv = batch_normalization(x)
    b. conv = Conv2D(x)
    c. return conv
2) def first(x):
    a. conv = Conv2D(x)
    b. conv = conv_block(conv)
    c. sht = Conv2D (x)
    d. sht = batch_normalization(sht)
    e. res = Add()([conv, sht])
    f. return res
3) def residual(x):
    a. res = conv_block(x)
    b. res = conv_block(res)
    c. shortcut = Conv2D (x)
    d. shortcut = batch_normalization (shortcut)
    e. output = Add()([shortcut, res])
    f. return output
4) def up_concat(x, direct):
    a. up = UpSampling2D((2, 2))(x) # upscales rows and columns
    b. up = Concatenate()([up, direct]) # upscales channels
    c. return up

5) def down(x, channels):
    a. x0, x1, x2, x3 = x.shape
    b. dwconv_3 = SeparableConv2D (x, dilation = 2)
    c. dwconv_3 = BatchNormalization(dwconv_3)
    d. dwconv_7 = SeparableConv2D (x, dilation = 4)
    e. dwconv_7 = BatchNormalization(dwconv_7)
    f. paddings = tf.constant([[0, 0], [0, 1], [0, 1], [0, 0]])
    g. if dwconv_3.shape[1] != int(x1 / 2) or dwconv_7.shape[1] != int(x1 / 2):
        i. if dwconv_3.shape[1] > dwconv_7.shape[1]:
            1. dwconv_3 = dwconv_3[:, 0:-1, 0:-1, :]
            2. dwconv_7 = tf.pad(dwconv_7, paddings, "REFLECT")
    h. elif dwconv_7.shape[1] > dwconv_3.shape[1]:
        i. dwconv_3 = tf.pad(dwconv_3, paddings, "REFLECT")
        ii. dwconv_7 = dwconv_7[:, 0:-1, 0:-1, :]
    i. if dwconv_3.shape[1] == int(x1 / 2) and dwconv_7.shape[1] == int(x1 / 2):
        i. dwconv_3 = tf.pad(dwconv_3, paddings, "REFLECT")
        ii. dwconv_7 = tf.pad(dwconv_7, paddings, "REFLECT")
    j. f = Concatenate()([dwconv_3, dwconv_7])
    k. return f
6) def spatial(x, dilation=1):
    a. x1, x2, x3 = x.shape[1], x.shape[2], x.shape[1]
    b. x = down(x, x3)
    c. x = SeparableConv2D(filters=x3 * 2, kernel_size=2, dilation_rate=dilation)(x)
    d. x = BatchNormalization()(x)

7) def gate(x, channels, reduction_ratio=16):
    a. g = Dense(channels // reduction_ratio, use_bias=False)(x)
    b. g = BatchNormalization()(g)
    c. g = Activation('relu')(g)
    d. g = Dense(channels, use_bias=False)(g)
    e. return g

8) def channel(x, reduction_ratio=16):
    a. x1, x2, x3 = x.shape[1], x.shape[3], x.shape[2]
    b. x = tf.reshape(Flatten()(x), [-1, x1, x2, x3])
    c. x_avg = tf.reshape(Flatten()(GlobalAvgPool2D()(x)), [-1, 1, x1, 1])
    d. x_max = tf.reshape(Flatten()(GlobalMaxPool2D()(x)), [-1, 1, x1, 1])
    e. x = tf.keras.layers.concatenate([x_avg, x_max], axis=1)
    f. x = Conv2D(1, kernel_size=(2, 1))(x)
    g. x = Flatten()(x)
    h. x = gate(x, x1, reduction_ratio)
    i. x = tf.reshape(Flatten()(x), [-1, x1, 1, 1])
    j. return x

9) def attn_mech(x, f):
    a. att_c = channel(x, reduction_ratio=16)
    b. att_s = spatial(x)
    c. w = sigmoid(tf.multiply(att_c, att_s))
    d. w = Conv2D(f, (1, 1))(w)
    e. return w

10) def resunet(input_size=(1024, 1024, 1)):
    a. f = 64
    b. inputs = Input(input_size)
    c. fir = first(inputs, f)
    d. dr1 = residual(fir, f * 2, strides=2)
    e. dr2 = residual(dr1, f * 4, strides=2)
    f. dr3 = residual(dr2, f * 8, strides=2)
    g. dr4 = residual(dr3, f * 16, strides=2)
    h. m1 = conv_block(dr4, f * 16, strides=1)
    i. m2 = conv_block(m1, f * 16, strides=1)
    j. ur1 = up_concat(m2, dr3)
    k. ur1 = attn_mech(ur1, f * 8)
    l. ur1 = tf.multiply(dr3, ur1)
    m. ur1 = residual(ur1, f * 16)
    n. ur2 = up_concat(ur1, dr2)
    o. ur2 = attn_mech(ur2, f * 4)
    p. ur2 = tf.multiply(dr2, ur2)
    q. ur2 = residual(ur2, f * 8)
    r. ur3 = up_concat(ur2, dr1)
    s. ur3 =attn_mech(ur3, f * 2)

t. ur3 = tf.multiply(dr1, ur3)
u. ur3 = residual(ur3, f * 4)
v. ur4 = up_concat(ur3, fir)
w. ur4 = attn_mech(ur4, f)
x. ur4 = tf.multiply(fir, ur4)
y. ur4 = residual(ur4, f * 2)
z. output = Conv2D(1, (1, 1), padding="same", activation="sigmoid")(ur4)
aa. model = Model(inputs, output)
bb. model.compile(optimizer=adam, loss=BinaryFocalLoss(gamma=2), metrics=[dice_coef])
cc. return model

11) Train model

<u>Flood_detection.py</u>

1) model = load(Attn_res_Unet)
2) ip1 = SAR img before flood
3) ip2 = SAR img after flood
4) op1 = model.predict(ip1)
5) op2 = model.predict(ip2)
6) flood_region = op1 XOR op2
7) area = correlate flood_region to area in sq. km
8) print(area)

5. Results and Discussion

5.1 Output

The model outputs a segmentation mask given a Satellite Aperture Image. The output consists of two classes, black and white which corresponds to land and water regions respectively. We have used two most prominent loss functions to find the most optimal one and found that, and these loss functions are Binary Cross Entropy and Focal Loss. We found that Focal loss has higher Intersection over Union score and dice score. The outputs using these 2 loss functions are shared below. The output when using Binary cross entropy as the loss function results in a speckled output with irregularities in the mask. These speckles are not robust and result in a poor quality mask. Whereas the output using Focal loss provides a much better output quality with decreased speckles and irregularities. These are crucial considering the use case of detecting floods, because every irregularity of white regions results in incorrect flood region detection resulting in a wrong approximation of flood area.

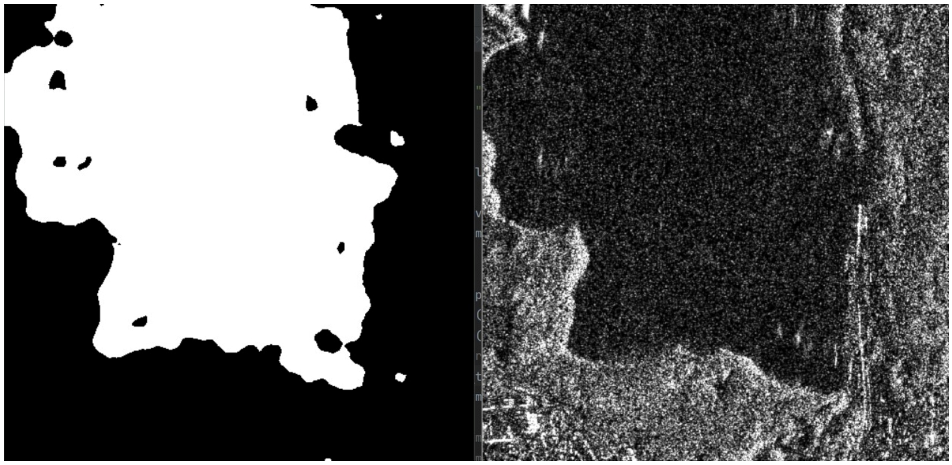


FIGURE 7.1. OUTPUT SAR MASK **FIGURE 7.2** INPUT SAR IMAGE

FIGURE 7. INPUT AND OUTPUT IMAGE USING FOCAL LOSS FUNCTION

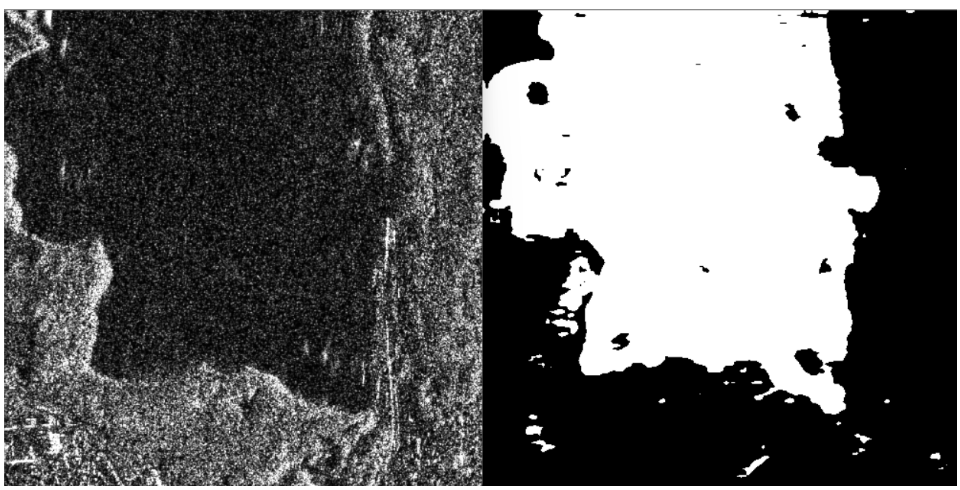


FIGURE 8.1. INPUT SAR IMAGE **FIGURE 8.2** OUTPUT SAR MASK

FIGURE 8. INPUT AND OUTPUT IMAGE USING BINARY CROSS-ENTROPY LOSS FUNCTION'

Once these masks are found, the flood regions is calculated. The proposal initial takes asinput, two satellite aperture radar images of the same geolocation before and after floods. The used model then finds the segmentation masks for these two satellite aperture radar images.In these masks, the white regions denote water and black regions denote the land. In order to find the affected land region, the after flood image is subtracted by the before flood image. The resulting white regions denote all those regions where the water has increased. Reading the documentation of the dataset, we found that 1 square pixel in the image, is equivalent to around
0.4 square kilometres. Multiplying the number of white pixels in the image, we get an

approximateregion in real world, which is affected by the flood.

FIGURE 9. FLOOD REGION

```
FLOOD AFFECTED AREA = 1877.2 sq. km
```

FIGURE 9. FLOOD AREA

5.2. Comparison with Existing models

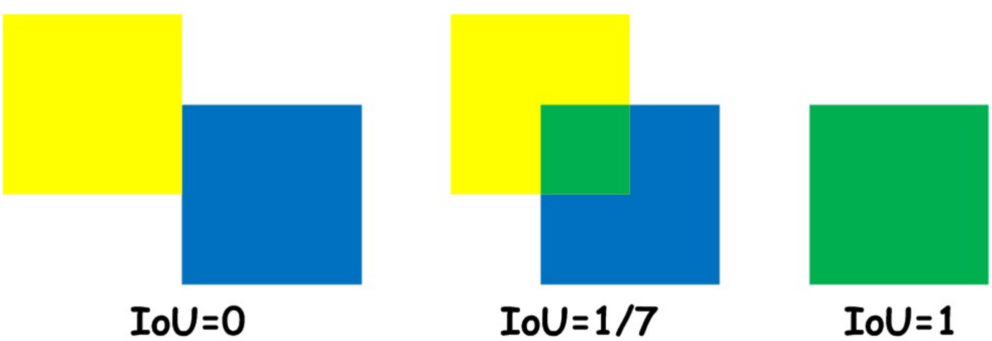
We compare our model with different approaches of SAR image segmentation.The overall accuracy of path vector is 9.44% which is lower among all the approaches and our model named u-net based attention network has the highest accuracy of 95%.

Approaches	Accuracy%
Patch Vector	9.44
Textures	75.03
Gabor Filters	65.17
BoW	87.28
SAE	85.36
MLFP	91.10
Unet-Attention	95.00

Table 1. Comparing Overall Accuracy of Different Model of Segmentation

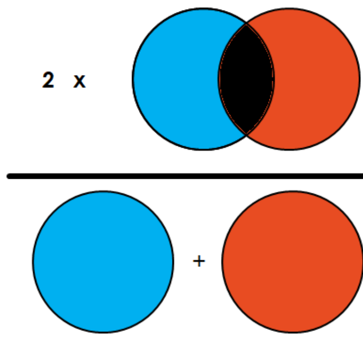
5.2.1. IOU Coefficient

To compute IoUs, the predictions are accumulated in a confusion matrix, weighted by sample_weight and the metric is then calculated from it. As indicated in the graphic to the bottom, the IoU is the area of overlap between the predicted segmentation and the ground truth divided by the area of union between the predicted segmentation and the ground truth. Our proposal received an IOU score of 91% for testing data of around 200 images.



5.2.2. Dice Coefficient

The dice similarity coefficient is a reproducibility validation metric and a geographic overlap index. Fleiss referred to it as the fraction of explicit agreement. A DSC's value ranges from 0 to 1, with 0 indicating no spatial overlap and 1 representing total overlap between two sets of binary segmentation results. The dice coefficient score is visualized below. Our proposal received an dice coefficient score of 94% for testing data of around 200 images.



5.3. Summary of the Result

The input to this model is SAR images. It has a very dense nature and has a high spatial resolution which can't be analysed using human eyes, Each input image of size 1024X1024 pix.And the output is a segmented image that will indicate the land and water covered area. The water is indicated by white regions and the land with black regions. This mask is then followed by subtracting SAR images before and after the flooding to calculate the flood region area in square kilometres

6. Conclusions

From our proposal, we conclude that segmenting the SAR images is crucial to get clarity on changes in the topography, waterbody, agriculture etc. There are various image segmentation algorithms but not all of them are suitable for SAR images because of their dense nature. Wealso conclude from our literature review that the attention unit is very beneficial for segmenting SAR images, which reduces training time and helps the model learn only the most important features from the upper layer's spatial information. Therefore, in this proposal we have developed an attention-based U-Net architecture where the vanilla convolution blocks are replaced by Residual Convolution blocks to extract finer features from the SAR image. Along with this, we will compare 4 different Attention unit architectures in this proposal and then choose the best one against otherprominent SAR image segmentation strategies. In the current proposal we have only being able to segment a SAR image into land and water. In future we also want to be able to segment various sub-features of lan such as agriculture, building, etc. This way, we will be able to pin point which region is most severely affected and also be able to estimate total losses. In the future we want to improve the performance of our architecture by increasing dataset size from 2000 to at-least 50000. This will guarantee better quality of masks and higher metric scores. Along with this, we also would like to train our model on an even more powerful system, capable of handling huge datasets.

The lateral resolution of the image obtained by numerical reconstruction was assessed utilizing a wavelet image decomposition and image correlation. The best lateral resolution obtained witha high NA recording, 164 nm, represents an improvement of more than a factor two relative to previously published results.

7. References

1. Wang, P., Zhang, X. and Hao, Y., 2019. A Method Combining CNN and ELM for Feature Extraction and Classification of SAR Image. *Journal of Sensors*, 2019, pp.1-8.
2. Zhang, L., Lu, S., Xiang, D. and Su, Y., 2021. SAR Image Ship Detection Based on a Superpixel Non-window Fast CFAR Strategy. *2021 SAR in Big Data Era (BIGSARDATA)*,.
3. Chen, S., Wang, H., Xu, F. and Jin, Y., 2016. Target Classification Using the Deep Convolutional Networks for SAR Images. *IEEE Transactions on Geoscience and Remote Sensing*, 54(8), pp.4806-4817.
4. Coman, C. and Thaens, R., 2018. A Deep Learning SAR Target Classification Experiment on MSTAR Dataset. *2018 19th International Radar Symposium (IRS)*,.
5. Shang, R., Wang, J., Jiao, L., Stolkin, R., Hou, B. and Li, Y., 2018. SAR Targets Classification Based on Deep Memory Convolution Neural Networks and Transfer Parameters. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(8), pp.2834-2846.
6. Zhang, T., Zhang, X., Ke, X., Liu, C., Xu, X., Zhan, X., Wang, C., Ahmad, I., Zhou, Y., Pan, D., Li, J., Su, H., Shi, J. and Wei, S., 2022. HOG-ShipCLSNet: A Novel Deep Learning Network With HOG Feature Fusion for SAR Ship Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 60, pp.1-22.
7. Rostami, M., Kolouri, S., Eaton, E. and Kim, K., 2019. Deep Transfer Learning for Few-Shot SAR Image Classification. *Remote Sensing*, 11(11), p.1374.
8. Pan, J., Wang, X. and Cheng, Y., 2016. Single-Sample Face Recognition Based on LPP Feature Transfer. *IEEE Access*, 4, pp.2873-2884.
9. Rostami, M., Kolouri, S., Eaton, E. and Kim, K., 2019. SAR Image Classification Using Few-Shot Cross-Domain Transfer Learning. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*,.
10. Zhong, C., Mu, X., He, X., Wang, J. and Zhu, M., 2019. SAR Target Image Classification Based on Transfer Learning and Model Compression. *IEEE Geoscience and Remote Sensing Letters*, 16(3), pp.412-416.
11. Zhang, F., Wang, Y., Ni, J., Zhou, Y. and Hu, W., 2020. SAR Target Small Sample Recognition Based on CNN Cascaded Features and AdaBoost Rotation Forest. *IEEE Geoscience and Remote Sensing Letters*, 17(6), pp.1008-1012.
12. Shahzad, M., Maurer, M., Fraundorfer, F., Wang, Y. and Zhu, X., 2019. Buildings Detection in VHR SAR Images Using Fully Convolution Neural Networks. *IEEE Transactions on Geoscience and Remote Sensing*, 57(2), pp.1100-1116.
13. Emek, R. and Demir, N., 2020. BUILDING DETECTION FROM SAR IMAGES USING UNET DEEP LEARNING METHOD. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, XLIV-4/W3-2020, pp.215-218.
14. Pai, M., Mehrotra, V., Verma, U., & Pai, R. (2020). Improved Semantic Segmentation of Water Bodies and Land in SAR Images Using Generative Adversarial Networks. *International Journal Of Semantic Computing*, *14*(01), 55-69. doi: 10.1142/s1793351x2040003
15. …….
16. Zhang, Z., Yang, J., & Du, Y. (2022). Deep Convolutional Generative Adversarial Network With Autoencoder for Semisupervised SAR Image Classification. *IEEE Geoscience And Remote Sensing Letters*, *19*, 1-5. doi: 10.1109/lgrs.2020.3018186
17. Rostami, M., Kolouri, S., Eaton, E., & Kim, K. (2019). Deep Transfer Learning for Few-Shot SAR Image Classification. *Remote Sensing*, *11*(11), 1374. doi: 10.3390/rs11111374
18. Bai, X., Xue, R., Wang, L., & Zhou, F. (2019). Sequence SAR Image Classification Based on Bidirectional Convolution-Recurrent Network. *IEEE Transactions On Geoscience And Remote Sensing*, *57*(11), 9223-9235. doi: 10.1109/tgrs.2019.2925636
19. Sharifzadeh, F., Akbarizadeh, G., & Seifi Kavian, Y. (2018). Ship Classification in SAR Images Using a New Hybrid CNN–MLP Classifier. *Journal Of The Indian Society Of Remote Sensing*, *47*(4), 551-562. doi: 10.1007/s12524-018-0891-y
20. Pant, T., Singh, D., & Srivastava, T. (2010). Advanced fractal approach for unsupervised classification of SAR images. *Advances In Space Research*, *45*(11), 1338-1349. doi: 10.1016/j.asr.2010.01.008

21. Zhao, Z., Jiao, L., Zhao, J., Gu, J., & Zhao, J. (2017). Discriminant deep belief network for high-resolution SAR image classification. *Pattern Recognition*, *61*, 686-701. doi: 10.1016/j.patcog.2016.05.028
22. Zhang, A., Yang, X., Fang, S., & Ai, J. (2020). Region level SAR image classification using deep features and spatial constraints. *ISPRS Journal Of Photogrammetry And Remote Sensing*, *163*, 36-48. doi: 10.1016/j.isprsjprs.2020.03.001
23. Geng, J., Jiang, W., & Deng, X. (2020). Multi-scale deep feature learning network with bilateral filtering for SAR image classification. *ISPRS Journal Of Photogrammetry And Remote Sensing*, *167*, 201-213. doi: 10.1016/j.isprsjprs.2020.07.007
24. Wang, Z., Fu, X., & Xia, K. (2022). Target Classification for Single-Channel SAR Images Based on Transfer Learning With Subaperture Decomposition. *IEEE Geoscience And Remote Sensing Letters*, *19*, 1-5. doi: 10.1109/lgrs.2020.3027363
25. Chen, S., Wang, H., Xu, F., & Jin, Y. (2016). Target Classification Using the Deep Convolutional Networks for SAR Images. *IEEE Transactions On Geoscience And Remote Sensing*, *54*(8), 4806-4817. doi: 10.1109/tgrs.2016.2551720
26. Zhu, X., Montazeri, S., Ali, M., Hua, Y., Wang, Y., & Mou, L. et al. (2021). Deep Learning Meets SAR: Concepts, models, pitfalls, and perspectives. *IEEE Geoscience And Remote Sensing Magazine*, *9*(4), 143-172. doi: 10.1109/mgrs.2020.3046356
27. Rostami, M., Kolouri, S., Eaton, E., & Kim, K. (2019). SAR Image Classification Using Few-Shot Cross-Domain Transfer Learning. *2019 IEEE/CVF Conference On Computer Vision And Pattern Recognition Workshops (CVPRW)*. doi: 10.1109/cvprw.2019.00120
28. Zhang, Y., Sun, X., Sun, H., Zhang, Z., Diao, W., & Fu, K. (2018). High Resolution SAR Image Classification with Deeper Convolutional Neural Network. *IGARSS 2018 - 2018 IEEE International Geoscience And Remote Sensing Symposium*. doi: 10.1109/igarss.2018.8518829
29. Liu, Z., Lai, R., & Guan, J. (2022). Spatial and Transform Domain CNN for SAR Image Despeckling. *IEEE Geoscience And Remote Sensing Letters*, *19*, 1-5. doi: 10.1109/lgrs.2020.3022804
30. Santhi, V., Mohandass, D., Jayanthi, J., Arulmozhivarman, P., & Mehra, R. (2021). Speckle Reduction in SAR Images using CNN. *2021 3Rd International Conference On Signal Processing And Communication (ICPSC)*. doi: 10.1109/icspc51351.2021.9451809
31. Dera, D., Rasool, G., Bouaynaya, N., Eichen, A., Shanko, S., Cammerata, J., & Arnold, S. (2020). Bayes-SAR Net: Robust SAR Image Classification with Uncertainty Estimation Using Bayesian Convolutional Neural Network. *2020 IEEE International Radar Conference (RADAR)*. doi: 10.1109/radar42522.2020.9114737
32. Li, Y., Li, X., Sun, Q., & Dong, Q. (2022). SAR Image Classification Using CNN Embeddings and Metric Learning. *IEEE Geoscience And Remote Sensing Letters*, *19*, 1-5. doi: 10.1109/lgrs.2020.3022435
33. Wu, Z., Hou, B., & Jiao, L. (2021). Multiscale CNN With Autoencoder Regularization Joint Contextual Attention Network for SAR Image Classification. *IEEE Transactions On Geoscience And Remote Sensing*, *59*(2), 1200-1213. doi: 10.1109/tgrs.2020.3004911
34. Fu, K., Zhang, T., Zhang, Y., Wang, Z., & Sun, X. (2022). Few-Shot SAR Target Classification via Metalearning. *IEEE Transactions On Geoscience And Remote Sensing*, *60*, 1-14. doi: 10.1109/tgrs.2021.3058249
35. Zhao, Z., Li, J., Luo, Z., Li, J., & Chen, C. (2021). Remote Sensing Image Scene Classification Based on an Enhanced Attention Module. *IEEE Geoscience And Remote Sensing Letters*, *18*(11), 1926-1930. doi: 10.1109/lgrs.2020.3011405
36. Jing, H., Sun, X., Wang, Z., Chen, K., Diao, W., & Fu, K. (2021). Fine Building Segmentation in High-Resolution SAR Images Via Selective Pyramid Dilated Network. *IEEE Journal Of Selected Topics In Applied Earth Observations And Remote Sensing*, *14*, 6608-6623. doi: 10.1109/jstars.2021.3076085
37. Zhao, Z., Jia, M., & Wang, L. (2021). High-Resolution SAR Image Classification via Multiscale Local Fisher Patterns. *IEEE Transactions On Geoscience And Remote Sensing*, *59*(12), 10161-10178. doi: 10.1109/tgrs.2020.3039847
38. Sharma, S., Senzaki, K., & Aoki, H. (2018). Comparative Study of Feature Extraction Approaches for Ship Classification in Moderate-Resolution SAR Imagery. *IGARSS 2018 - 2018 IEEE International Geoscience And Remote Sensing Symposium*. doi: 10.1109/igarss.2018.8518966

39. Zhao, J., Guo, W., Liu, B., Zhang, Z., Yu, W., & Cui, S. (2017). Preliminary exploration of SAR image land cover classification with noisy labels. *2017 IEEE International Geoscience And Remote Sensing Symposium (IGARSS)*. doi: 10.1109/igarss.2017.8127696
40. Wang, C., Gu, H., & Su, W. (2022). SAR Image Classification Using Contrastive Learning and Pseudo-Labels With Limited Data. *IEEE Geoscience And Remote Sensing Letters*, *19*, 1-5. doi: 10.1109/lgrs.2021.3069224
41. Wang, J., Ren, Y., & Wei, S. (2019). Synthetic Aperture Radar Images Target Detection and Recognition with Multiscale Feature Extraction and Fusion Based on Convolutional Neural Networks. *2019 IEEE International Conference On Signal, Information And Data Processing (ICSIDP)*. doi: 10.1109/icsidp47821.2019.9172989
42. Tanase, R., Datcu, M., & Raducanu, D. (2016). A convolutional deep belief network for polarimetric SAR data feature extraction. *2016 IEEE International Geoscience And Remote Sensing Symposium (IGARSS)*. doi: 10.1109/igarss.2016.7730968
43. Zhou, Y., Jiang, X., Li, Z., & Liu, X. (2020). SAR Target Classification with Limited Data via Data Driven Active Learning. *IGARSS 2020 - 2020 IEEE International Geoscience And Remote Sensing Symposium*. doi: 10.1109/igarss39084.2020.9324364
44. Ismail, A., Gao, X., & Deng, C. (2014). SAR image classification based on texture feature fusion. *2014 IEEE China Summit &Amp; International Conference On Signal And Information Processing (Chinasip)*. doi: 10.1109/chinasip.2014.6889221
45. He, C., Liu, X., Han, G., Kang, C., & Chen, Y. (2017). Fusion of statistical and learnt features for SAR images classification. *2017 IEEE International Geoscience And Remote Sensing Symposium (IGARSS)*. doi: 10.1109/igarss.2017.8128247
46. Emek, R., & Demir, N. (2020). BUILDING DETECTION FROM SAR IMAGES USING UNET DEEP LEARNING METHOD. *The International Archives Of The Photogrammetry, Remote Sensing And Spatial Information Sciences*, *XLIV-4/W3-2020*, 215-218. doi: 10.5194/isprs-archives-xliv-4-w3-2020-215-2020