ABSTRACT

Forest fire is becoming one of the most significant natural disasters at the expense of ecology and economy. In this article, we develop an effective Squeeze Net based asymmetric encoder-decoder U-shape architecture, Attention U-Net and SqueezeNet (ATT Squeeze U-Net), mainly functions as an extractor and a discriminator of forest _re. This model takes attention mechanism to highlight useful features and suppress irrelevant contents by embedding Attention Gate (AG) units in the skip connection of U-shape structure. In this way, salient features are emphasized so that the proposed method could be competent at forest fire segmentation tasks with a small number of parameters. Specifically, we first replace classical convolution layer by a depthwise one and engage a Channel Shuf_e operation as a feature communicator in the Fire module of classical SqueezeNet. Then, this modified SqueezeNet is employed as a substitution of the encoder of Attention U-Net and a corresponding DeFire module designed is combined into the decoder as well. Finally, to classify true _re, we take use of a fragment of the encoder in ATT Squeeze U-Net. The experimental results of modified SqueezeNet integrated Attention U-Net show that a competitive accuracy at 0.93 and an average prediction time at 0.89 second per image are achieved for reliable real-time forest fire detection.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ABBREVIATION	EXPANSION

CHAPTER 1
INTRODUCTION

1.1 Introduction

Early detection and identi_cation of forest _re can avoid damaging disaster. Fire detection methods such as satellite-based detection, optical sensing, wireless sensing and remote sensing gain notable improvements to forest _re alarm. In this study, we focus on monitoring _re detection driven by computer vision. Computer vision mechanisms for _re detection could be mainly classi_ed into two categories, traditional image processing method and deep Convolutional Neural Network (CNN) method. Existing conventional detection algorithms mainly operate based on visual properties of _re, such as color, spectral, texture, motion and geometric features. Despite the low cost and simplicity, traditional methods strongly rely on appropriate feature description of _re. Some natural phenomena, such as sunset and fog would cause false alarm and missing report to these approaches occasionally. To solve these problems, a more advanced _re detection scheme proposing the use of CNN technology instead of feature description has attracted more and more attention. Meanwhile, recent development of GPU allows the use of CNN-based methods for _ame detection. Common disadvantage of them seems to be that large datasets are required to learn the best features. As a result of this, the model would over-_tted under huge training dataset, whereas applying a small number of dataset for learning would be insuf_cient. Recently, some lightweight compression networks [1] that could achieve real-time processing have been introduced when reasonable mistakes are allowed. In this article, we propose an ef_cient neural network architecture for forest _re detection and recognition based on Attention U-Net and SqueezeNet (ATT Squeeze U-Net). The proposed framework consists of two stages, a segmentation module extracting the shape of forest _re, and a classi_cation module identifying whether the detected _re area is true or not.We remove the encoder of conventional Attention U-Net [2] and place our modi_ed SqueezeNet with new designed Fire modules at the contraction path. Besides, some implementations are carried out on traditional Fire module [3] in SqueezeNet. We replace the original 3 _ 3 convolution layer with a Depthwise Convolution (DWConv) kernel, and a Channel Shuf_e operation is added to have feature communication enhanced. We design corresponding DeFire modules for ATT Squeeze U-Net model and embed them into the decoder for better up-sampling. We then develop a new classi _cation framework for _re identi_cation by reusing a part of the encoder of ATT Squeeze U-Net. A discussion of how many output feature maps of the encoder layers are chosen is raised to reach a most effective selection for subsequent _re recognition. We evaluated the ATT Squeeze U-Net on some publicly available datasets, and the results demonstrate that the proposed framework can produce better _re area extraction results than some existing algorithms.

1.2 Future Scope

It can be summarizing that the future scope of the project circles around maintaining information regarding: It helps us to determine the efficient and more accurate results for recognition of handwritten digits. Various challenges are identified which may provide more lively interest to the researchers. The main challenge is to identify the diverse human writing styles, different angles, different shapes and size of digits, pure input quality, low accuracy rate in recognition etc. Hence, a lot of research work is done to solve these problems.Handwritten Digit Recognition is the capacity of a computer to interpret the manually written digits from various sources like Messages, Bank cheques ,Papers and pictures, Postal mail sorting, Form data entry , Digits entered in any forms etc.

CHAPTER 2
LITERATURE SURVEY

2.1 Literature Survey

Fire attack in forest makes major degradation in the forest environment and ecosystems. Forest fire detection in the early stage can prevent major causes due to fire attack. A novel digital image processing based on multi-feature best decision-based forest fire detection (MF-BD-FFD)is proposed in this work. To increase the sensitivity of detection, color and texture feature with hybrid decision making algorithms such as artificial neural networks (ANN), Support Vector machine (SVM), k- nearest classifier (KNN) is used and optimized output will be selected. By using proposed method, the accuracy of the system is increased with a factor of 5% when compared to the conventional technique.

The Convolutional Neural Network (CNN) has been used in many fields and has achieved remarkable results, such as image classification, face detection, and speech recognition. Compared to GPU (graphics processing unit) and ASIC, a FPGA (field programmable gate array)-based CNN accelerator has great advantages due to its low power consumption and reconfigurableproperty. However, FPGA's extremely limited resources and CNN's huge amount of parameters andcomputational complexity pose great challenges to the design. Based on the ZYNQ heterogeneousplatform and the coordination of resource and bandwidth issues with the roofline model, the CNN accelerator we designed can accelerate both standard convolution and depthwise separable convolution with a high hardware resource rate. The accelerator can handle network layers of different scales through parameter configuration and maximizes bandwidth and achieves full pipelined by using a data stream interface and ping-pong on-chip cache. The experimental results show that the accelerator designed in this paper can achieve 17.11GOPS for 32bit floating point when it can also accelerate depthwise separable convolution, which has obvious advantages compared with other designs.

We present a novel and practical deep fully convolutional neural network architecture for semantic pixel-wise segmentation termed SegNet. This core trainable segmentation engine consists of an encoder network, a corresponding decoder network followed by a pixel-wise classification layer. The architecture of the encoder network is topologically identical to the 13 convolutional layers in the VGG16 network [1]. The role of the decoder network is to map the low resolution encoder feature maps to full input resolution feature maps for pixel-wise classification. The novelty of SegNet lies is in the manner in which the decoder upsamples its lower resolution input feature map(s). Specifically, the decoder uses pooling indices computed in the max-pooling step of the corresponding encoder to perform non-linear upsampling. This eliminates the need for learning to upsample. The upsampled maps are sparse and are then convolved with trainable filters to produce dense feature maps. We compare our proposed architecture with the widely adopted FCN [2] and also with the well known DeepLab-LargeFOV [3], DeconvNet [4] architectures. This comparison reveals the memory versus accuracy trade-off involved in achieving good segmentation performance. SegNet was primarily motivated by scene understanding applications. Hence, it is designed to be efficient both in terms of memory and

computational time during inference. It is also significantly smaller in the number of trainable parameters than other competing architectures and can be trained end-to-end using stochastic gradient descent. We also performed a controlled benchmark of SegNet and other architectures on both road scenes and SUN RGB-D indoor scene segmentation tasks. These quantitative assessments show that SegNet provides good performance with competitive nference time and most efficient inference memory-wise as compared to other architectures. We also provide a Caffe implementation of SegNet and a web demo.