5G Rural Integrated Testbed

D3.13 Interim Final Report

Agricultural Use Cases

Acronym:	**5GRIT**
Full Title:	5G Rural Integrated Testbed
Coordinator:	Cybermoor
Project Duration:	04/2018 - 03/2019
Website:	www.5GRIT.co.uk

Partner:	Precision Decisions
Contact:	Stephen Leese
Email:	stephen@precisiondecisions.co.uk
Website	www.precisiondecisions.co.uk
Work Package:	WP3
Deliverable:	D3.13
Title:	Final Report
Lead:	Stephen Leese

Authors: Stephen Leese, Precision Decisions Ltd
Hamideh Kerdegari, Manzoor Razaak; Kingston University

Internal Review: Rachel Hardy, CYB

I

Table of Contents

1 Executive Summary

Technology development in rural areas has huge potential - possibly even higher than in urban areas due to the clarity of the use cases. Funded by the Department for Digital, Culture, Media and Sport 5G Testbeds & Trials Programme, the 5G Rural Integrated Testbed project (5GRIT) and this paper set out and describe these use cases to illustrate the technology used, to assess the benefits, and to recommend future steps.

Uses cases which at first sight seem mundane have high value when we include the value of the opportunity created through the time saved. Especially in remote, extensive farming, direct farm income often has to be supplemented by non-farm income. The technologies applied here create such an opportunity by saving the farmer time, which can then be used to generate supplementary income.

Two uses cases were assessed in detail:

- monitoring of sheep in extensive hill-land areas;
- monitoring of arable crops in intensive farming areas.

Economic benefits were assessed and although these are challenging to evaluate, we estimate an economic value of almost £400 for a sheep farm and up to £92 per hectare additional margin for arable crops.

The technology is still very young and recommended next steps would include further refinement of the specific algorithms to monitor eiter crops or animals.

2 Introduction

As part of the Department for Digital, Culture, Media and Sport 5G Testbeds & Trials Programme, this paper reports on the deliverables and experience of the 5GRIT project to assess use cases for a 5G infrastructure in rural areas. It includes descriptions of the use cases based on personas determined at the start of the project, work completed and progress made as well as feedback from future potential users of an appropriate service based on a 5G network.

The 5G network is not a simple, linear extension of 3G and 4G. It does provide a step change in the speed and volume of data which can be transmitted from rural areas. This ability provides opportunities to create use cases based on data transmission in rural areas which so far, with 3G and 4G, have not been possible.

Such use cases are potentially more beneficial in rural areas than urban use cases for 5G, since in cities, there is already an excellent fibre optic network which allows for transmission of vast amounts of data, for example, live streaming of surgical procedures between hospitals.

As the project included a partner which specialises in drone capabilities, we looked at the use of drones to enhance and illustrate the use cases. There are other methodologies of capturing video images of agricultural operations, such as video cameras mounted on tractor booms, and these are mentioned here but have not been part of a specified use case. The benefit from such technology is, however, very comparable with the images captured by drones.

5GRIT project aims to develop a platform based on 5G enabled UAV for automatic sustainable crop and livestock monitoring. Computer vision technology is an integral part of the 5GRIT project for developing intelligent systems that enable smart farming applications such as crop-weed classification and livestock monitoring. The objective of the computer vision algorithms are to facilitate crop monitoring to identify weeds from crops and livestock monitoring with capabilities to identify and track livestock including livestock counting, health monitoring and anomaly detection.

An illustration of the aerial crop/livestock monitoring system is presented in Figure 1. The system includes capturing images of the farm to provide as an input data for the deep learning based computer vision algorithms that are designed to detect presence of weed/ livestock. The algorithms process the images at a data processing centre and transmit the outcomes to the concerned stakeholders, such as a farmer.

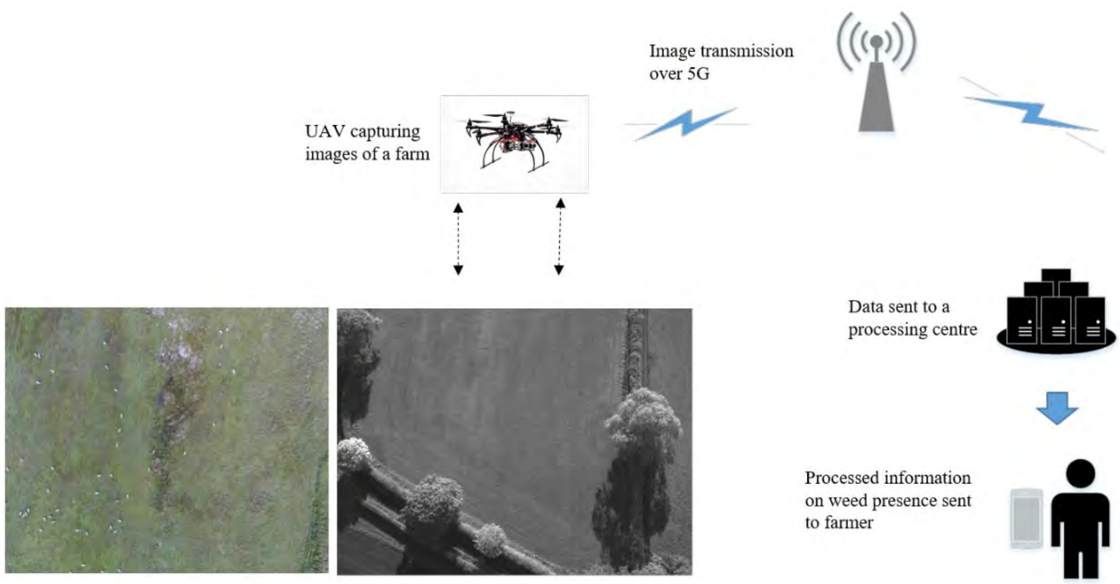

Figure 1. A weed/livestock detection system comprising of a UAV with a mounted multispectral/RGB camera, 5G transmission system, and a data processing centre hosting the captured image data and weed detection algorithms.

In this document, we report the implementation of the deep learning algorithms for crop/weed classification, detection and counting of livestock during the project. In Section 3, the research question addressed by the project is discussed. Section 4 presents the implementation of computer vision algorithms applied for two use cases and also discusses the crowdsourcing of image annotations using Amazon Mechanical Turk platform. Results and learning points are discussed in section 5. The final sections cover security, economic benefits of the use cases, and future work and conclusion.

3 Research Question - what we set out to achieve

Can 5G deliver productivity improvements through smart agriculture for upland livestock farmers and lowland arable farmers through improved monitoring and analysis of data gathered by drones and analysed in the cloud?

To answer this question, we derived personas of potential participants (see Appendix 1) and from these, derived agricultural operations we estimated would show most improvement from the use of an enhanced data streaming via a 5G network.

An important goal for 5GRIT project was to demonstrate the capabilities of 5G for smart farming applications. The project envisaged a data acquisition and transmission system mounted on a drone to capture imagery of the farms and transmit the farm images over 5G for artificial intelligence based computer vision analysis.

At the onset of the project, two broad use cases were selected: arable use case and livestock use case. The research aim was to develop and implement computer vision algorithms to assist farmers in decision-making related to the use cases. To outline the research objectives, apart from extensive literature review, discussions were held with farmers to understand their expectations from the technology. Based on the discussions and research, the following objectives were set out:

- ➢ provide a farm remote monitoring system to assist farmers to allow surveying of large farms with relative ease and high speed
- ➢ use drone imagery to identify anomaly areas in the farm
- ➢ enable large area crop growth monitoring in a short time
- ➢ aerially monitor livestock over large farms for identification, verification and distribution of livestock presence

Based on the above objectives and feasibility analysis, we set out to research and develop algorithms primarily designed to perform three common farming related operations:

1. ***Weed-crop classification***: capture imagery data of the farms and then analyse the farms to identify areas of weed growth on the farms
2. ***Livestock counting***: to identify and count the livestock presence in the farms
3. ***Plant counting***: monitor plant growth and yield via automated plant counting algorithm

4 Implementation

The use cases were split into two main areas, arable and livestock.

Arable: For the arable use case, we focused on management of in-field variability in order to enable farmers to move from a prophylactic approach to crop protection and move to a more targeted approach to pest control. To do so required the capturing of images at multiple times during a full cropping season, which for winter wheat in the UK is from September through to the following August.

Photos were taken using drones and the images processed by Kingston University.

Farmers were then asked whether this type of information would enable them to make better management decisions, resulting in a more optimized use of crop protection inputs which in turn leads to higher yields and especially, higher margins.

Livestock: the use case here focused on a key activity *hill-land sheep farmers* undertake on a daily basis and which takes up a lot of their time. Large and extensive farms can be hard to manage with vast amounts of land to be covered to track animal movement. This means animals can be hard to locate, and it can be difficult to spot abnormalities in animals' behaviours.

Often, especially in the summer months, hill sheep graze on fell pastures which can be a number of miles from the farmhouse. These fells are also quite extensive and the sheep have a large area to roam. Farmers like to know on a daily basis where their animals are and whether they need help if sick or injured. To do this, farmers go out daily to monitor their animals. These are counted and observed and this entire operation can take up to three hours per day depending on the number of animals and the distance of the fell from the farmhouse.

The use case was derived to assess if the use of a drone, streaming live images back to an office where, following appropriate image processing, an algorithm will estimate the number of sheep on the fell and give an indication to the farmer of any potential sick or injured animals. These often separate themselves from the main group of animals so can be detected.

From the technical perspective, artificial intelligence or deep learning methods such as *convolutional neural networks* typically require large sets of training images from which they can identify object-specific features (object being a plant or weed or crop) and "learn" to develop the object detection ability. The creation of a large corpus of image datasets is a common challenge during the data acquisition process. Availability of large datasets would lead to development of more robust and accurate object identification and classification algorithms.

The image data acquired for 5GRIT mainly includes a combi of multispectral and RGB images for the crop-weed classification use case and standard RGB for the livestock monitoring use case. The lack of large training dataset is a resource limitation for the algorithms in 5GRIT. To overcome the challenge of limited amount of training datasets, a class of deep learning algorithms named Generative Adversarial Networks (GANs) are utilized for both use cases. The GAN system can generate photorealistic images that appear authentic to human observers with features that are closer to real images. Thus, in circumstances that lack large amounts of input images for training the algorithms, GAN can be used to generate images that are similar to the original images and increase the image corpus for training. The images acquired in the 5GRIT are given as inputs to the GAN system and it generates photorealistic images that are similar to the images. Therefore, the challenge of lack of large datasets for training algorithms is overcome with the help of the

GAN method. Using the generated images by GAN, the proposed deep learning algorithms are applied to disambiguate crop from weed and count livestocks.

4.1 Arable Use Case 1: Crop/weed classification

The computer vision algorithm for the crop-weed classification use case employs deep learning methods for automated analysis of the images captured via unmanned aerial vehicles. Deep learning methods require very large corpus of data for training and hence to generate sufficient amount of data for algorithm training, Generative Adversarial Networks (GANs) [1] system are explored to generate additional corpus of training images for algorithm training and development. The *semi-supervised GAN* method that is explored for the 5GRIT algorithm training, especially focusing on the multispectral data of the crop-weed classification use case.

4.1.1 Semi-supervised GAN for Crop/Weed Classification

The GAN framework was first introduced by Goodfellow et al. [1] to train deep generative models. A GAN usually contains two networks: a generative (G) network and a discriminative (D) network. Both networks G and D are trained simultaneously in an adversarial manner, where G tries to generate fake inputs as real as possible, and D tries to disambiguate between real and fake data. Unlike typical GAN, where the discriminator is a binary classifier for discriminating real and fake images, semi-supervised GAN implements a multiclass classifier. In semi-supervised learning, where class labels are not available for all training images, it is convenient to leverage unlabeled data for estimating a proper prior to be used by a classifier for enhancing performance.

We extend typical GAN by replacing the traditional discriminator D with a fully convolutional multiclass classifier, which, instead of predicting whether a sample x belongs to the data distribution (real or fake), it assigns to each input image pixel a label y from the n classes (i.e. crop, weed or background) or mark it as a fake sample (extra n + 1 class). More specifically, D network predicts the confidence for n classes of image pixels and softmax is employed to obtain the probability of sample x belonging to each class. Figure 2 presents a schematic description of the semi-supervised GAN architecture that three inputs such as generated multispectral data, unlabeled data and a small number of labeled data are fed into the discriminator. Note that our GAN formulation is different from typical GANs, where the discriminator is a binary classifier for discriminating real/fake images, while our discriminator performs multiclass pixel categorization.

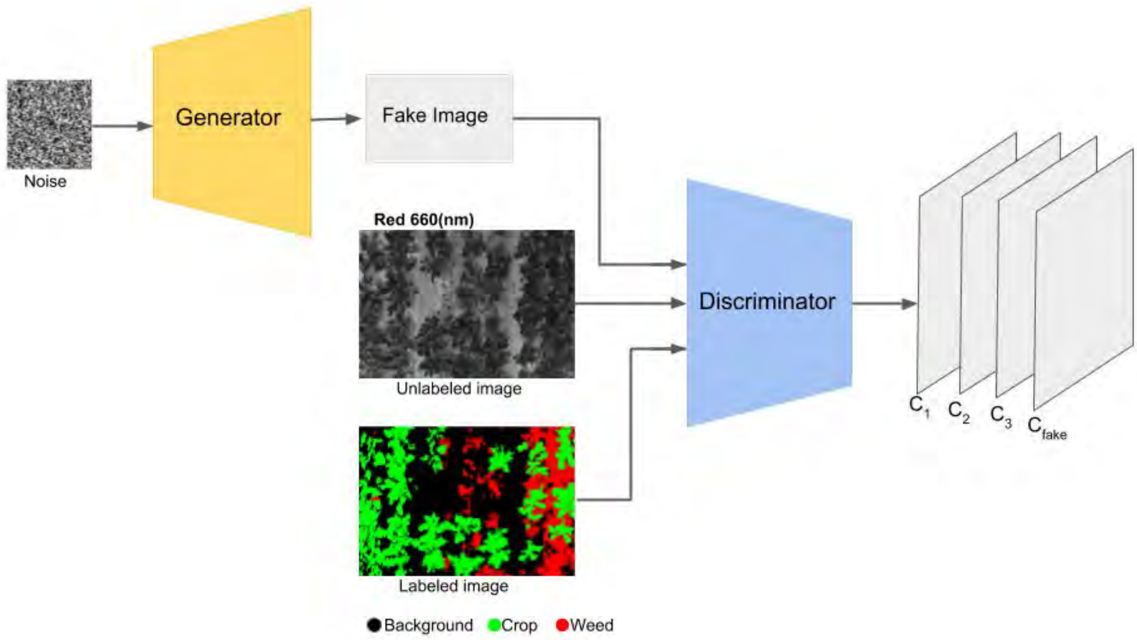

Figure 2: The semi-supervised GAN architecture. Random noise is used by the Generator to generate an image. The Discriminator uses generated data, unlabeled data and labeled data to learn class confidences and produces confidence maps for each class as well as a label for a fake data.

Figure 3 shows some generated images by the generator from different channels. These images indicate that the semi-supervised GAN framework is able to learn spatial object patterns, for example, crop shape and weed shape and generate images realistic to the original ones.

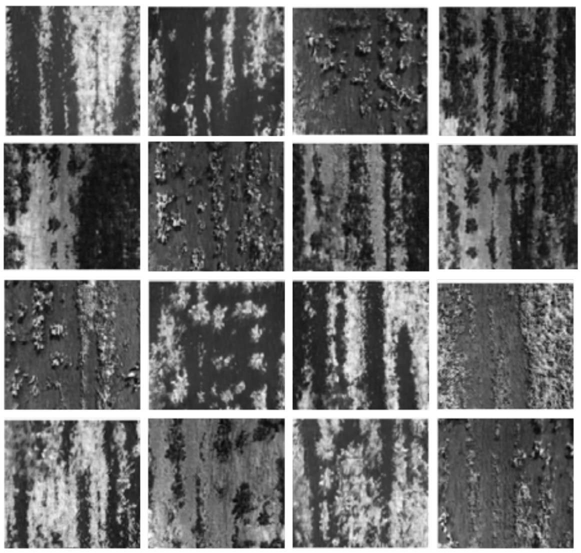

Figure 3. Images generated by the generator of the semi-supervised GAN on the weedNet dataset. Interestingly, patterns related to crops and weeds from NDVI, Red and NIR channel can be observed that highlights the effectiveness of the approach.

4.1.2 System Overview

The details of our semi-supervised GAN architecture including both generator and discriminator are presented in Figure 4. The generator network takes a uniform noise distribution as input, followed by a series of four convolution layers and generates a fake image resembling samples from real data distribution. The discriminator network processes the generated images, unlabeled images and a small number of labeled multispectral images to learn class confidence, producing a confidence map for each class as well as a label for fake data. The underlying idea is that adding large fake multispectral images forces real samples to be close in the feature space, which, in turn, improves classification accuracy. Our semi-supervised GAN formulation extends the canonical GAN, where the discriminator is a binary classifier for discriminating real/fake images, implementing a pixel-wise multiclass classifier.

Note that the proposed semi-supervised GAN is adopting DCGAN [2] architecture with a modification in the last layer of the discriminator by replacing sigmoid activation function with softmax to enable pixel-wise multiclass classification. All the networks are implemented using Keras library with Tensorflow backend. The standard Adam optimizer with momentum is used for the discriminator and the generator optimization with learning rate and β1 (momentum) set to 0.0002 and 0.5, respectively. A batch size of 32 and batch normalization are utilized for both networks. ReLU activation function is applied in the generator for all layers except for the output, which uses Tanh and LeakyReLU activation in the discriminator for all layers. In the experiments, no data augmentation or post-processing is performed.

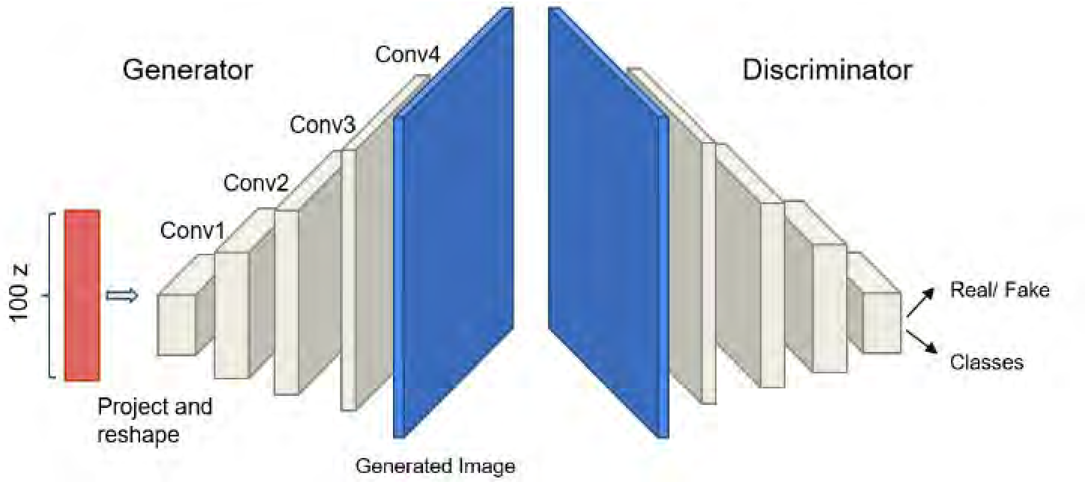

Figure 4: The network architecture of our semi-supervised GAN. The noise is a vector of size 100 sampled from a uniform distribution and is used as input to the generator. The number of feature maps in the four different convolutional layers, respectively, are 256, 128, 64, 32 and 1 (Here 1 shows the number of channels).

During the testing process, the discriminator network is only used as pixel-wise multiclass classifier network. Given a test image, the softmax layer of the discriminator outputs a set of probabilities of each pixel belonging to semantic classes, and accordingly, the label with the highest probability is assigned to the pixel.

4.1.3 Results

This section presents the experimental setup, followed by a quantitative and qualitative evaluation of the crop/weed classification use case. The proposed semi-supervised GAN algorithm is evaluated on the weedNet [3] dataset collected by a micro aerial vehicle (MAV) equipped with a 4-band Sequoia multispectral camera. The multispectral images are captured from sugar beets field at 2 meter height. The dataset contains only NIR and Red channel due to difficulties in image registration of other bands. From corresponding NIR and Red channel images, the Normalized NDVI, given by:1,000

$$NDVI = (NIR - Red) / (NIR + Red)$$

is extracted indicating the difference between soil and plant. Therefore, each training/test image consists of the 790nm NIR channel, the 660nm red channel, and NDVI imagery. The dataset contains only crops, or weeds, or crop-weed combination along with their corresponding pixel-wise annotated data.

For semi-supervised training, different percentages of pixel-wise annotated images (such as 50%, 40% and 30%) are used as labeled data to the discriminator and the rest of images are without pixel-wise annotations. As metric, F1 score measure that is a harmonic average of precision and recall is employed:

$$F1 = 2^* (precision * recall) / (precision + recall)$$

Where precision is TP/(TP + FP), recall is TP/(TP + FN), TP, FP and FN indicate the number of true positive, false positive and false negative, respectively. Quantitative results of our method on weedNet are shown in Table 1.

Table 1. Results on the weedNet dataset using 50%, 40% and 30% of labeled data with different number of channels. Higher F1 values indicate better classification performance.

F1 Score	*Amount of labeld data*					
	50%		40%		30%	
Channel	*Crop*	*Weed*	*Crop*	*Weed*	*Crop*	*Weed*
Red	0.831	0.814	0.822	0.813	0.792	0.813
NIR	0.839	0.823	0.80	0.821	0.782	0.773
NDVI	0.826	0.803	0.817	0.79	0.788	0.812
Red + NIR	**0.857**	**0.865**	0.837	0.834	0.823	0.815
Red + NIR + NDVI	0.852	0.831	0.847	0.821	0.816	0.812

F1 measure with a varying number of input channels and different amount of labelled data are used as evaluation metric in this experiment. Considering the difficulty of the dataset, all models (including different channels + different amount of labeled data) perform reasonably well (about 80% for all classes). As shown in Table 1, two input channels (Red and NIR) yield higher performance compare to single channels as they contain more useful features to be used by the semi-supervised GAN network. However, using 3 channels (NDVI + Red +

NIR) did not improve performance as NDVI depends on NIR and red channels rather than capturing new information.

Furthermore, the network was evaluated by reducing the amount of labeled data starting at 50% and then reducing by step 10 to 30% to find out how it affects the classification performance. It is expected that higher amount of labeled data result in better performance. It can be seen by comparing the results of the 50%, 40% and 30% in Table 1.

The qualitative evaluation is performed on four sample test images. As shown in Figure 5 , each row contains original Red channel, NIR channel, NDVI imagery, semi-supervised GAN probability output and the corresponding ground truth. The probability of each class is mapped to the red, green and black color representing weed, crop and background, respectively. There are some noticeable weed and crop misclassification areas in the images that occur mostly when crops and weeds are surrounded by each other. This misclassification shows that network can capture high-level features such as shape and texture in addition to the low-level features.

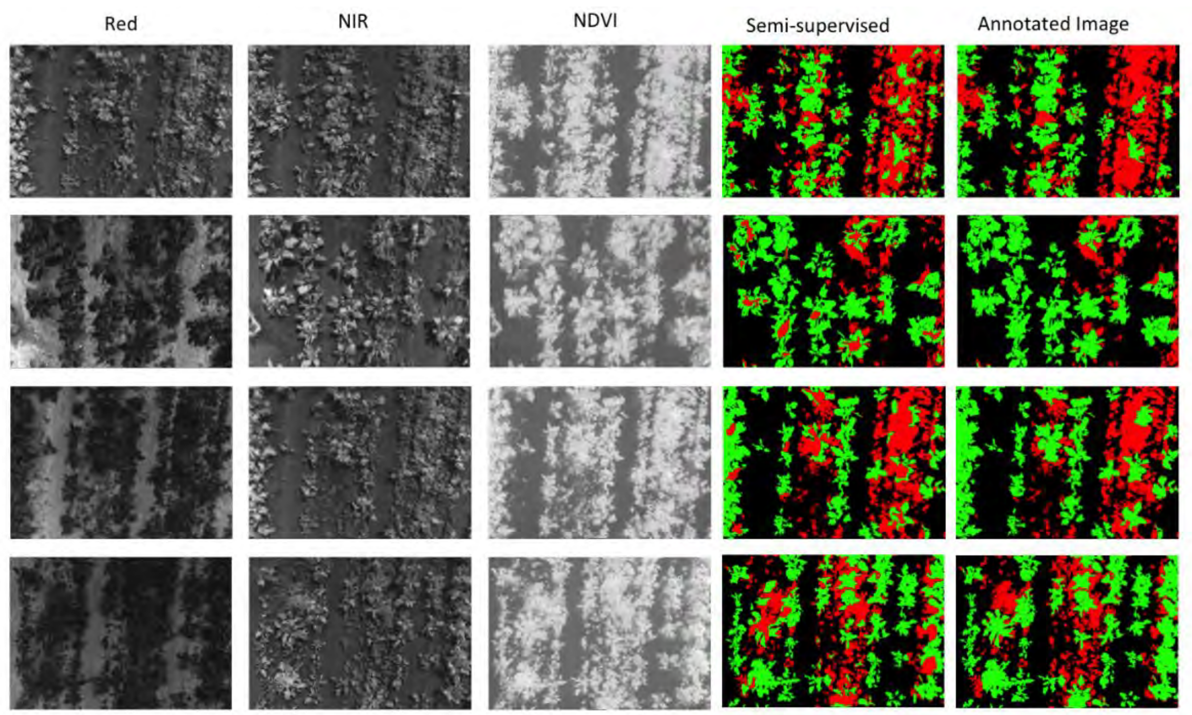

Figure 5. Qualitative results for the weedNet test set. The first three columns are input data to the semi-supervised GAN, the fourth is the results of semi-supervised GAN using 30% o1,000f labeled data and the last column is ground truth.

4.2 Arable Use Case 2: Plant Counting

The objective of use case 1 was to disambiguate plant from weed and it was explored using *weedNet* dataset as explained in the previous section. However, due to lack of enough weed in this season, it is not possible to apply this algorithm on the collected UAV data that contain only plants. Therefore, plant counting is suggested as an alternative objective for use case 1. This section describes a preliminary experiment that has been performed to count the plants in the collected UAV RGB imagery data.

4.2.1 Template Matching for Plant Counting

A new objective in the use case for crop monitoring is to develop algorithms that can fetch the plant count from images to allow farmers to estimate crop growth. As an extension of the previous algorithms, to develop the plant counting algorithm, template matching technique [4] was applied on the plant RGB data collected. Template matching is a technique in image processing that finds areas of an image that match (are similar) to a template image (patch). As can be seen in Figure 6, the first row is sample image patches that the algorithm uses to find their corresponding matches in the input images. The second row shows the results of the template matching method. It can be observed that the method is able to identify crop patches from the input images. The method also generates errors as shown in the middle image.

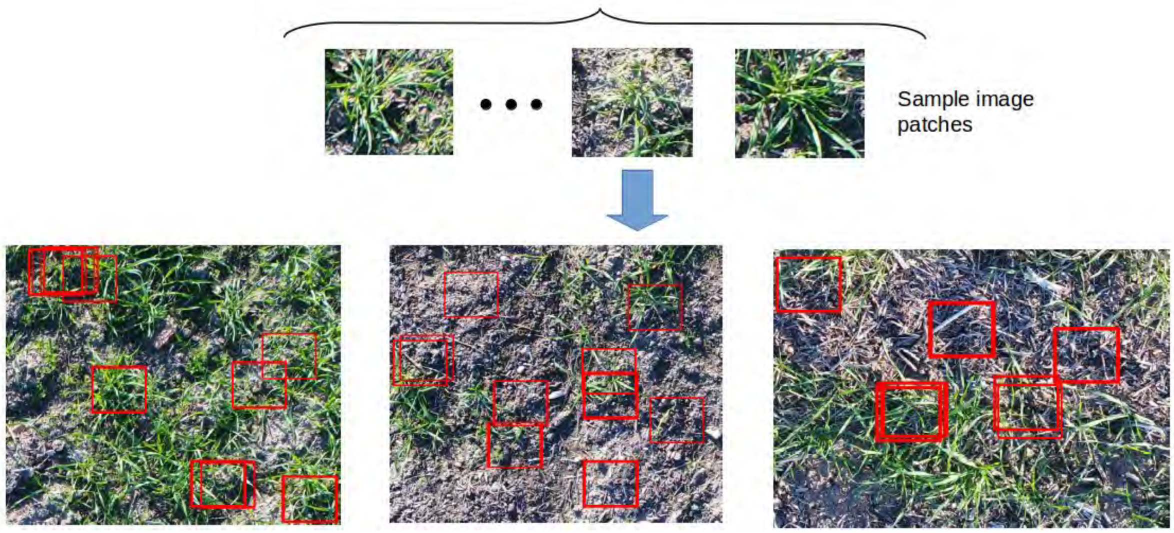

Figure 6: An example of applied template matching on some sample plant images.

This experiment with template matching method shows limited success in plant counting. More effective methods are being studied for detection and counting of plants, especially overlapping plants.

4.3 Livestock Use Case: Livestock monitoring

The use case aims at developing algorithms capable of livestock detection and counting from aerial images captured by drones. State-of-the-art object detection algorithms were adapted and implemented for the livestock detection algorithm. Aerial images of livestock from a drone flown at 50m altitude were captured by *Blue Bear Systems Research* at a farm in Alston. Flying the UAV at high altitudes enables us to capture larger field of view of the farms and consequently reduces flight duration required to cover the entire farm area.

At higher altitudes, the mounted cameras capture lower resolution images and the target such as livestock are represented in small sizes in the images that pose challenges for computer vision analysis. Figure 7 shows example images of livestock that are captured by the mounted cameras on drones at 50m altitude. It can be observed that target livestock (sheep) are represented in small sizes in the images. Generally, small size targets in images have lower feature representations and provides limited learning information for machine learning algorithms. Recent deep learning methods based on convolutional neural networks (CNN) such as Fast-CNN and *you only look once* (YOLO) have demonstrated high detection accuracy, however, they usually perform better with large targets than small targets in the images. In aerial based livestock monitoring, targets are represented in small resolution size and thus makes target detection more challenging. To address the livestock detection use case in 5GRIT, an algorithm based on image super-resolution and CNN based object detection method is developed and tested.

Figure 7. Images of livestock captured from a UAV at 50m altitude. Left column shows the original images. The right column images zoomed versions of the livestock targets. At high altitudes, livestock are captured as small target objects in the images.

The aim of the algorithm is to present a method that represents the low-resolution size livestock with better feature representations to facilitate easier target detection. The system studied in our tests utilizes the enhanced super-resolution GAN (ESRGAN) [5] method for generating super-resolution images from the low-resolution farm images. The generated super-resolution images provide better feature representations of the livestock and can enhance the improvements in target detection. The you only look once (YOLO) target object detection system is applied on the generated super-resolution images and the low-resolution images for livestock detection and for performance comparison of target objects, i.e. livestock.

4.3.1 System Overview

Figure 8 is a block representation of our implemented target detection system. The system consists of two main components: (a) the enhanced SRGAN model and (b) the YOLO object detection system. The ESRGAN model takes the low resolution image input to generate the corresponding 4X up-scaled high resolution images. The generated high-resolution images and the low-resolution images are then provided as inputs to the YOLO object detector for comparative performance of target detection between the two images. The ESRGAN and the YOLO system are briefly described below.

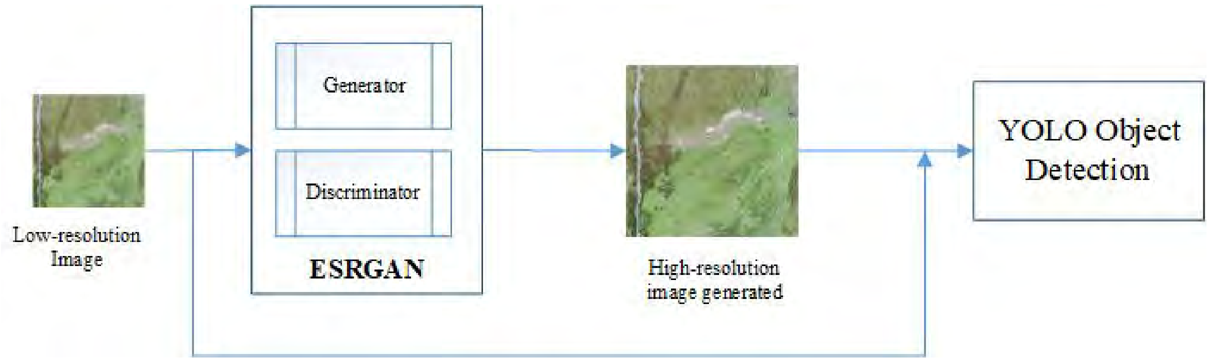

Figure 8. The ESRGAN and YOLO object detector based system used in our study to generate super-resolution images for small object detection on livestock images.

A. Enhanced Super-Resolution GAN

The ESRGAN is based on the seminal work of [6] on SRGAN with an aim to enhance the visual quality of the generated images. The ESRGAN introduces a Residual-in-Residual Dense Block (RRDB) to the generator network architecture for easier training and better image generation capacity of the network. The RRDB is based on residual learning methods that helps in reducing the number of layers required for a deep network to train. RRDB enables easier learning for the neural networks. The discriminator of the network is based on relativistic average GAN (RaGAN) that discriminates the relative realism of the generated image to the original image. Further, an enhanced perceptual loss function is proposed that leads to generation of super-resolution images with sharper edges and better visual quality.

B. you only look once (YOLO)

The you only look once (YOLO) is a widely popular state-of-the-art, target object detection system proposed in [7]. In comparison to other target detection methods, YOLO uses a simple approach. The object detection is considered as a single regression problem where a single convolutional network predicts multiple bounding boxes on an image and their associated class probabilities. Unlike other target detection methods that use sliding window approaches, YOLO considers the entire image during training, thus, it takes in the contextual information of the objects and classes to be detected. The YOLO method has shown high detection accuracy, at par with other state-of-the-art methods such as R-CNN.

As a first step in our system implementation, the ESRGAN is trained with low resolution UAV images of livestock to obtain corresponding high resolution images with potential better feature representations of the target small objects (in our case, sheep). In the second step, the generated super-resolution images from the ESRGAN are provided to the YOLO detection network for detecting presence of sheep in the drone images.

4.3.2 Results

The images generated by the ESRGAN network against the corresponding low resolution images are shown in Figure 9. It can be observed that the ESRGAN generates better resolution images when compared to the low-resolution images provided. The sheep in the low-resolution images (represented as white dots at 50m altitude) are blurry and have lower feature representations when compared to the super-resolution images generated by the ESRGAN. The average peak signal to noise ratio (PSNR) of the generated images were approximately 23 dB.

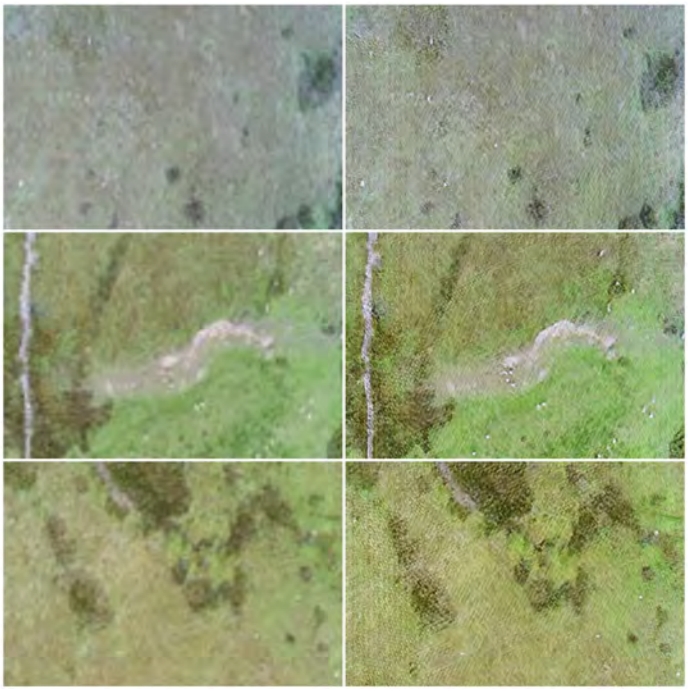

Figure 9. The low-resolution (left column) and the corresponding generated super-resolution representations (right column) by ESRGAN of livestock images

The visual results of the images show that ESRGAN can generate higher resolution of small targets from low-resolution images. The ESRGAN network, however, has limitations as the images consist of blocky artifacts. The results can be improved with potentially higher iterations and larger image dataset for training. Further, other super-resolution based GAN networks to be explored to generate better super-resolution images of small targets.

The generated images and the low-resolution images were used for target detection testing on the trained YOLO model. The livestock targets were detected at a precision of 57.3% by the YOLO model on the generated images. On the low-resolution images, the YOLO model was not able to perform livestock detection due to the low-resolution of the small livestock objects. The overall performance of YOLO on small object detection was not satisfactory. The results are consistent with the findings in the literature that have shown that YOLO detector performs poorly on small target detection [8][9][10].

The results of our study have shown that small target detection is a challenge and specifically for drone based images as the target objects are captured in small size. Super-resolution methods can provide high-resolution images, however, to achieve better target detection accuracy, different methods of target detection need to be explored and developed.

4.4 Amazon Mechanical Turk

A key aspect of implementation of the algorithms is availability of labelled images for training the algorithms to learn and perform the required tasks. Labelled images are annotated images where the presence of the target object such as plants and livestock are annotated with bounding boxes on the image. The location of the box on the image is used as data by the algorithms to identify the presence of the target object and learn its features in the image.

Annotating images is a time consuming and manually laborious process as it requires an individual to go through each image and identify the location of the target such as plant or weed or sheep and place a bounding box around it. To simplify this process, crowdsourcing model was used in our studies for image annotation.

Crowdsourcing model for image annotation is a process in which a large number of individuals are recruited over the internet to perform the image annotation. The individuals are briefed about the task requirements and are provided specific instructions, if any, on annotating the objects.

We required large database of labelled images to train our algorithms. Crowdsourcing approach was required to get the labelling. Amazon Inc. provides a crowdsourcing marketplace called Amazon Mechanical Turk (AMT) to outsource and recruit workers for image annotation. We used various Amazon web services such as cloud storage, mechanical turk and image annotation platform to crowdsource the image labelling. Figure 10 outlines the stages of crowdsourced image labelling process.

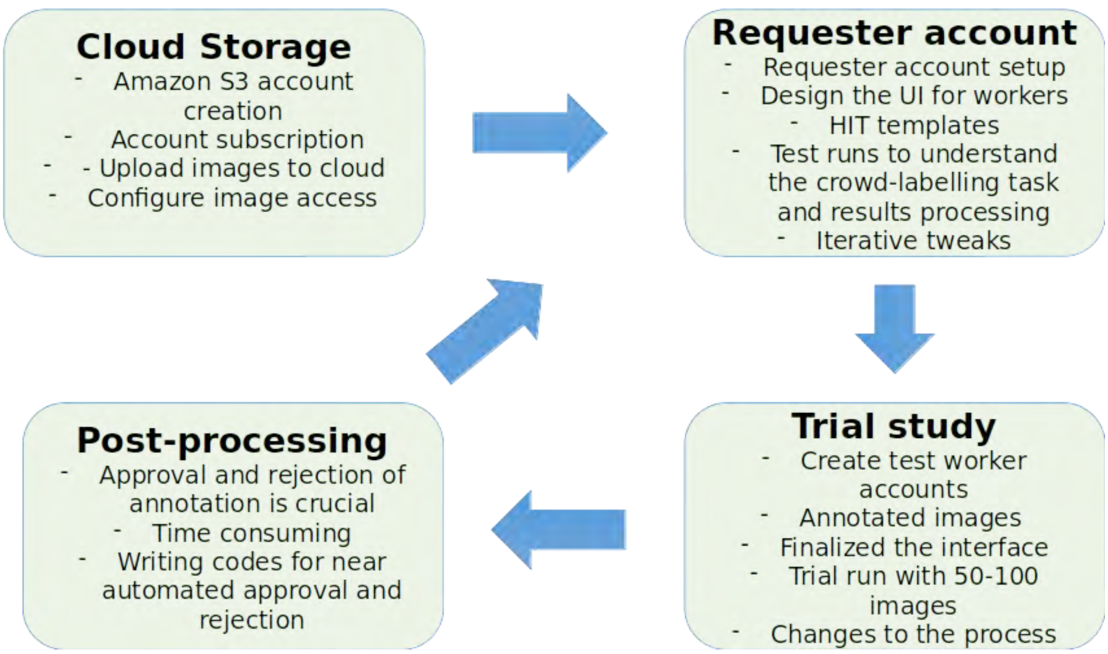

Figure 10. A block diagram outlining the crowdsourcing process for image annotation.

To summarise the crowdsourcing process on AMT, the images were uploaded to Amazon`s cloud storage. On the AMT platform, a UI was designed based on the template option provided where the workers can view the images and place bounding boxes on the target objects, i.e. plant and sheep. The instructions on the task requirements were provided along with an example annotated image for the workers. For our project, we uploaded around 4000 images for the plant classification and livestock monitoring use cases. Through AMT platform, we were able to get annotations for all the images in a matter of few hours. The annotations obtained were reviewed by us manually and then approved or rejected based on the quality of the annotations.

Below are few annotated images obtained from AMT platform. Both accepted and rejected annotations are shown. These annotated images were used for training our deep learning algorithms.

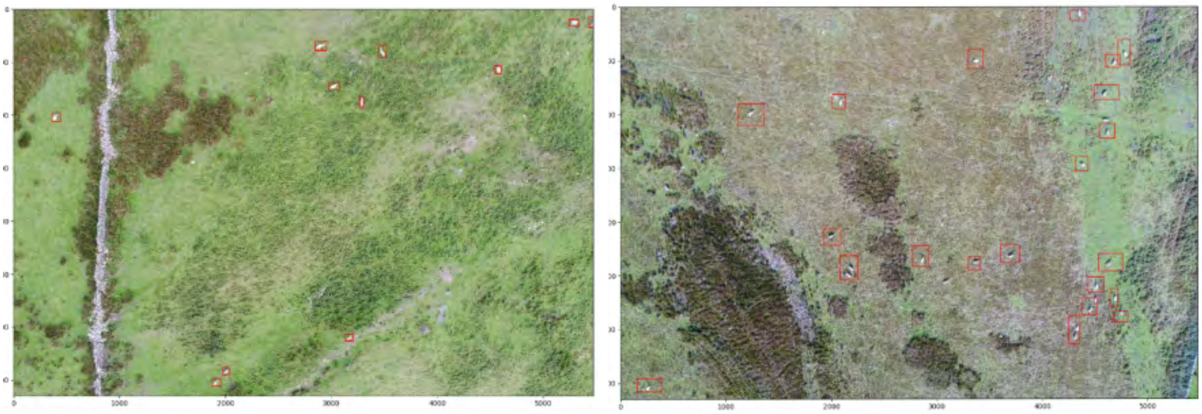

Figure 11. Approved annotated images with bounding boxes on livestock

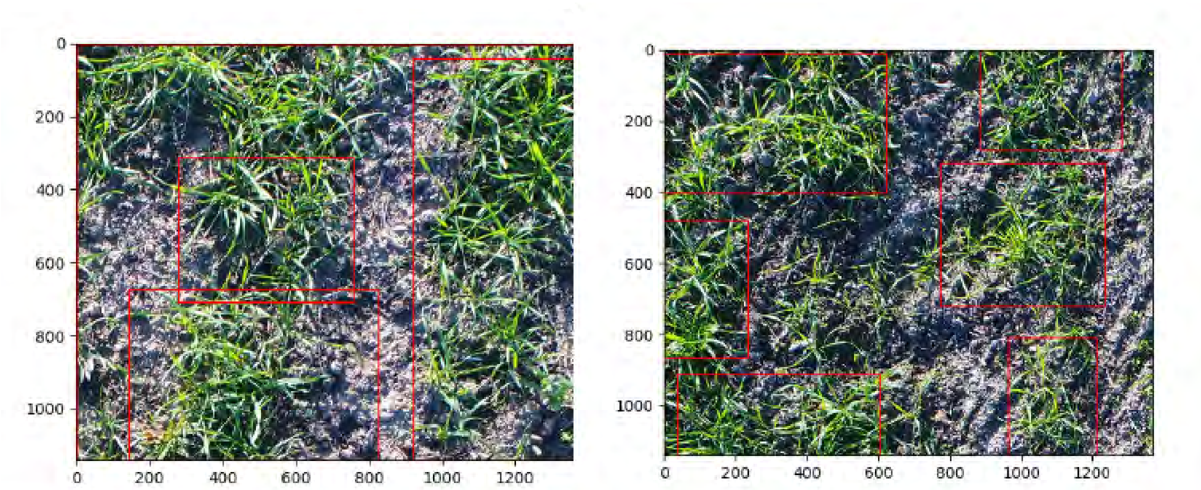

Figure 12. Rejected annotated images due to poor annotations by the workers

5 Key Learning Points

The data acquisition and algorithm development processes were intense and challenging mainly because unlike previous research studies presented in the literature on similar use cases, our study was not in a controlled environment. Real-practical scenarios poses obstacles were addressed during the trial studies. The key learning points from the project are:

- *Stakeholder enthusiasm* – Interactions with farmers while developing the use cases indicated there is a general enthusiasm towards adopting advanced technologies such as UAV and 5G for PA applications

- *Flight altitude* - Selecting an optimal altitude for flying UAV depends on factors such as area of the farm to be covered, uphills and slopes across the farm, resolution of the camera payload, feasible image quality at the altitude, and data requirements for the algorithm.

- *Data acquisition* – The data acquisition stage for the project did not coincide with the seeding season. For agricultural use case, acquiring data throughout the lifecycle of the crop would provide richer database to develop more deep-learning based applications.
- *Algorithm accuracy* – The deep-learning algorithms developed during the project demonstrated their capabilities for crop-weed classification and livestock detection with good accuracy. The performance of the algorithms could be further improved by acquiring more data at different cropping seasons and weather conditions to create more diverse dataset for the algorithms to learn from.

6 Security review

A physical security review was performed by Cybermoor to identify the security threats that exist on the project. Precision Decisions' response to the security review provided details that would be pertinent to use cases, such as vehicle theft from farming locations etc. but did not address issues that may be had within the company. We have not yet been directed to the company's security managers so will assume they do not have a specific security management entity (SME) in place. Perhaps as companies do not have a dedicated SME and have not had security incidents yet.

The recent integration of Precision Decision into the much larger Map of Ag, will introduce new points for consideration and this will be included in Phase 2.

Detailed security review

7 Economic Benefits

Animal Monitoring

The economic use case for this concerns the closer monitoring of the farm animals – in this case sheep – which enables cost benefits to farmers in two ways:

1. An overall reduction in medicine usage due to a targeted approach to animal health rather than a prophylactic approach. Current median spend is £6.82 per animal on a hill sheep farm. We estimate that a 5% reduction can be achieved through this targeted approach which results in a cost saving of £0.34 pence per animal on veterinary products. On a farm with 895 sheep (average size), this would equate to £0.34x895=£304 per year.
2. We think that as a result of better monitoring of animals grazing in remote areas, monitoring for disease and health issues (such as inflamed hooves) can be improved and issue identified earlier. In discussions with farmers, we estimate that 1 call per year could be saved as a result of monitoring the animals more closely. This equates to a saving of £50-£90 p.a. based on the current cost of a vet call out on these farms.

Overall, both savings could amount to £350 - £390 per year on the average sheep farm. While this is rather low, with today's very small margins of such farms, this is a sizeable contribution.

Arable crop production

The use case here is similar in that increased monitoring of crop health and diseases – particularly in real-time - will result in a lower application usage. Assuming the following as an example on winter wheat:

- Yield = 7.2 tons/hectare
- Price = £147/t
- Revenue = £1,058
- Cost of growing = £487/t
- Agrochem cost = £258/t
- Gross Margin = £571/t

We are estimating that through increased real-time monitoring of crops, we can reduce input of agrochemicals by 5%. This equates to an input reduction cost of 5% x £258 = £12.9/t yield. This in turn equates to a £92.8/ha saving or increase in margin as a result of closer crop monitoring.

8 Future Work

Future work includes adapting the algorithm for a near real-time application involving the transmission of aerial farm images including plants and livestock from UAV to a processing server over 5G wireless network.

For the plant/weed classification use case, it would be suggested testing the proposed semi-supervised GAN algorithm with other multispectral dataset that contain more channels such as Red edge, Green and Blue to investigate the effect of each channel separately and in combination with each other on the crop/weed classification performance.

In the plant counting use case, since the template matching algorithm does not detect the plants correctly due to overlapping plants as shown in Figure 4, Single Shot MultiBox Detector (SSD) [11] is suggested for future work that is more accurate and robust for object detection purpose. Further, acquiring image data of the crops at different stages of growth could be used to develop algorithms for crop growth monitoring applications.

Future work for the livestock monitoring includes exploring more advanced object detection methods such as Feature Fusion Single Shot Multibox Detector (FSSD) and develop a framework that is designed to perform more accurate object detection on UAV acquired images. It is also suggested to acquire more data from different altitudes to assess and identify an optimal flying altitude for livestock monitoring.

Further, a grid interface that includes an image mosaic of the farm captured from the UAV with a superimposed interface that indicates the distribution of the livestock over the farm can be an easy and quick visualization method for the farmers to monitor their livestock.

9 Conclusions

Intelligent methods for monitoring of crops and livestock can contribute towards improving precision farming applications. The 5GRIT project proposes using 5G enabled drones equipped with cameras to monitor farm lands with minimal human intervention in the field. This deliverable provides a report of WP3 related to the agricultural use cases aspect of the project. Under this work package, a project lifecycle of developing deep learning algorithms with a focus on two use cases; crop/weed classification and livestock monitoring, were developed.

For the crop/weed classification use case, a semi-supervised framework, based on Generative Adversarial Networks (GAN), for the classification of multi-spectral crop/weed images is proposed. The semi-supervisd GAN network is trained on the weedNet dataset captured by a micro aerial vehicle (MAV) from a sugar beet field. The results showed that a detection accuracy of 85% was achieved from the proposed method. Furthermore, the presented model generates synthetic images that could be used as additional multispectral data for other classifiers. For the livestock monitoring use case, an automated method of detecting the livestock from aerial images was developed. The algorithm adapted deep learning based object detection methods and super-resolution based GAN approach for achieving livestock monitoring from UAV images.

Future work includes extending the work on the algorithms to improve its performance accuracy. Several UAV flights could be conducted to acquire images of the crop at different growth stages and to conduct studies with adapted algorithms for near real-time applications.

References

[1] Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014. Generative adversarial nets. In *Advances in neural information processing systems* (pp. 2672-2680).

[2] Odena, A., 2016. Semi-supervised learning with generative adversarial networks. *arXiv preprint arXiv:1606.01583*.

[3] Sa, I., Chen, Z., Popović, M., Khanna, R., Liebisch, F., Nieto, J. and Siegwart, R., 2018. weedNet: Dense Semantic Weed Classification Using Multispectral Images and MAV for Smart Farming. *IEEE Robotics and Automation Letters*, *3*(1), pp.588-595.

[4] R. Brunelli, Template Matching Techniques in Computer Vision: Theory and Practice, *Wiley*, ISBN 978-0-470-51706-2, 2009.

[5] X. Wang, K. Yu, S. Wu, J. Gu, Y. Liu, C. Dong, C. C. Loy, Y. Qiao,and X. Tang, "ESRGAN: Enhanced super-resolution generative adversarial networks,"*arXiv preprint arXiv*:1809.00219, 2018.

[6] C. Ledig, L. Theis, F. Husz ´ar, J. Caballero, A. Cunningham, A. Acosta,A. P. Aitken, A. Tejani, J. Totz, Z. Wanget al., "Photo-realistic single image super-resolution using a generative adversarial network."in *CVPR*, vol. 2, no. 3, 2017, p. 4.

[7] YOLO," https://github.com/pjreddie/darknet.wiki.git, accessed: 2018-11-20.

[8] G. X. Hu, Z. Yang, L. Hu, L. Huang, and J. M. Han, "Small object detection with multiscale features,"*International Journal of Digital Multimedia Broadcasting*, vol. 2018, 2018.

[9] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, "Scalable object detection using deep neural networks," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp.2147–2154.

[10] Z. Ma, L. Yu, and A. B. Chan, "Small instance detection by integer programming on object density maps," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp.3689–3697

[11] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.Y. and Berg, A.C., "SSD: Single shot multibox detector", In *European conference on computer vision,* pp. 21-37, October, 2016.

Appendices

1 User Personas

The following user personas describe the profiles of farmers likely to use precision farming technology to identify a plant from weed, allowing them to count crops and increase productivity.

1.1 Head farmer, family farm

Name: Oscar
Age: 56
Gender: Male
Profession: Farmer
Marital Status: Married

Oscar is a farmer who has worked on his family farm in North Yorkshire since the age of 16. He inherited the farming operation from his Father as he did from his own Father. Oscar has two grown-up sons, neither of whom work the farm and a daughter who lives locally. The farm has always been family-run but he is finding that the younger generation are no longer interested in agriculture, his grandchildren being more likely to move away to the city to study or find work. As he is getting older, and with less workers available, he is looking for ways to simplify work and improve productivity. Oscar is not very good with technology, but does understand and appreciate its value. Oscar's daughter has an IT qualification and visits daily.

1.2 Farm Manager

Name: Rose
Age: 32
Gender: Female
Profession: Farm Manager
Marital Status: Single

Rose recently found work as a farm manager on Green Lane Farm after working a few years as assistant manager. Rose graduated with a BSc in Agriculture and Farm management from Newcastle University. Growing up in the countryside, she has been interested in agriculture since a young age and completed a farming apprenticeship when she left school. On joining the GLF team, Rose was keen to apply what she'd learnt about farm budgeting to increase productivity and sell a higher quality of crop. With farming subsidies falling, Rose feels that more of the budget should be spent on the automation of certain jobs, but is a little unsure about which system to use and the compatibility between them.

1.3 Farmer

Name: Tim
Age: 33
Gender: Male
Profession: Farmer
Marital Status: Single

Personal: Tim is 33 and supports his father with the family cattle farm just outside Lincoln. He is single and is interested in taking over the farm when his dad retires in the next 10 years. He is interested in new technologies and studied agriculture at Harper Adams University. He is interested in the farm becoming more efficient but struggles to persuade his dad and grandad that new technology can save them money. He is also concerned about connectivity between the machines and the farm office. He is aware of the benefits of new cloud technologies, but again is unsure about connectivity in a rural area where cell signal strength can be sadly lacking.

Type of persona

The type of persona can strongly influence the willingness to adopt new technology.

Map of Agriculture has vast experience with assessing these person types and has developed the following schema.

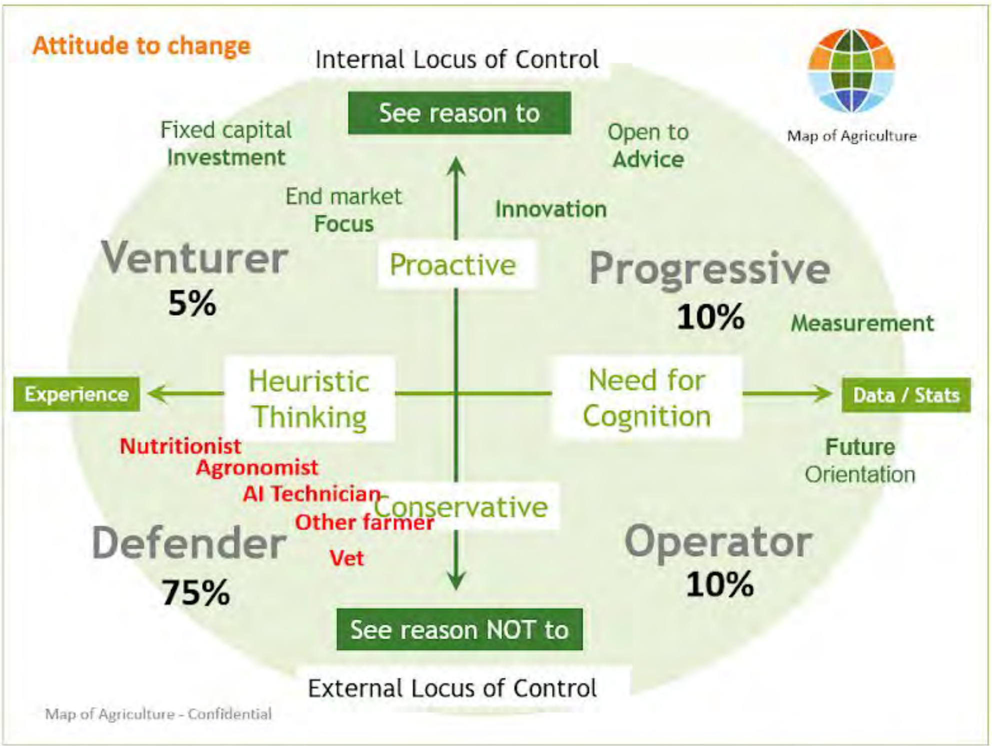

Figure 1: Attitude to change, Map of Agriculture.[1] ©

To identify the end users of each use case, we have taken into account data collected by Map of Agriculture, shown in figure 1, focusing on 'progressive' farmers. These farmers are considered to be innovative as well as open to new ideas and advice on farming practices.

This information is also useful in identifying ways to approach farmers according to their attitude toward change/technology.

[1] www.mapofagriculture.com