

2019

Multi-Spectral Visual Crop Assessment Under Limited Data Constraints

Patricia O'Byrne

Technological University Dublin, patricia.obyrne@tudublin.ie

Patrick Jackman

Technological University Dublin, patrick.jackman@tudublin.ie

Damon Berry

Technological University Dublin, damon.berry@tudublin.ie

Hector-Hugo Franco-Penya

Technological University Dublin, hector.franco@tudublin.ie

Michael French

Tanco Autowrap Ltd., mfrench@itanco.com

Follow this and additional works at: <https://arrow.tudublin.ie/scschcomcon>



Part of the [Artificial Intelligence and Robotics Commons](#)

See next page for additional authors

Recommended Citation

O'Byrne, P., Jackman, P., Berry, D., Franco-Penya, H.-H., French, M., and Ross, R. J. (2019). Multi-spectral visual crop assessment under limited data constraints. *Irish Machine Vision and Image Processing Conference 2019*, Grangegorman Campus, TU Dublin, 28 - 30 August .

This Conference Paper is brought to you for free and open access by the School of Computing at ARROW@TU Dublin. It has been accepted for inclusion in Conference papers by an authorized administrator of ARROW@TU Dublin. For more information, please contact yvonne.desmond@tudublin.ie, arrow.admin@tudublin.ie, brian.widdis@tudublin.ie.



This work is licensed under a [Creative Commons Attribution-Noncommercial-Share Alike 3.0 License](#)

Authors

Patricia O'Byrne, Patrick Jackman, Damon Berry, Hector-Hugo Franco-Penya, Michael French, and Robert J. Ross

Multi-spectral visual crop assessment under limited data constraints

Patricia O’Byrne¹, Patrick Jackman¹, Damon Berry¹, Hector-Hugo Franco-Penya¹, Michael French², and Robert J. Ross¹

¹*Technological University Dublin*

²*Tanco Autowrap*

Abstract

In an era of climate change and global population growth, deep learning based multi-spectral imaging has the potential to significantly assist in production management across a wide range of agricultural and food production domains. A key challenge however in applying state-of-the-art methods is that they, unlike classical hand crafted methods, are usually thought of as being only useful when significant amounts of data are available. In this paper we investigate this hypothesis by examining the performance of state-of-the-art deep learning methods when applied to a restricted data set that is not easily bootstrapped through pre-trained image processing networks. We demonstrate that significant result improvement can be obtained from deep residual networks over a baseline image processing model – even in the case where data collection is highly expensive and pre-trained networks cannot be easily built upon. Our work also constitutes a useful contribution to understanding the benefit of applying deep image multi-spectral processing techniques to the agri-food domain.

Keywords: Proximal Sensing, Deep Learning, Convolutional Neural Network.

1 Introduction

Precision agriculture has become increasingly relevant in the 21st century, when food production faces a new set of challenges, from population growth [UN, 2015] and diminishing resources [Gerland et al., 2014, Kummu et al., 2016]. Given these challenges, systems that can assist farmers in assessing the quality and state of crops can be highly valuable in optimising the production process. As a specific example, consider the case of milk production. Over 150 million households worldwide keep animals for milk [Kyte and Kaldani, 2013], and, ideally, grass silage should provide 20% to 30% per cent of an animal’s feed¹. Thus, when harvesting for silage, farmers benefit from knowing the biomass and moisture content of their pasture [Kung Jr, 2010, Tan et al., 2014]. Fortunately, bands of electromagnetic reflectance correlate to vegetation characteristics and hand crafted functions of reflected radiation – called vegetation indices – have been used for many decades to assist farmers in making such determinations [Xue and Su, 2017]. Unfortunately, these correlations vary according to plant phenology, ecology and sensing environment; meaning that the resultant hand-crafted functions are blunt tools that fail to take advantage of the true appearance of vegetation. We argue that methods based on full image processing and classification techniques, but that take into account a wide variety of spectral bands, have considerable potential to improve the estimation of crop quality.

While the application of deep image processing techniques has great potential benefit for a range of food producers, there are several complications that limit the direct application of these state-of-the-art methods. The most significant issue centres on the challenge of high-quality data collection. Data collection in agricultural and non-lab settings is expensive due to the overall challenge of sensor deployment, data collection,

¹<https://www.teagasc.ie/media/website/publications/2016/Teagasc-Quality-Grass-Silage-Guide.pdf>

and the costs of establishing accurate gold standard values for data. Due to these challenges the quantity of data collected in such settings is limited relative to the amounts of data typically expected for 'big data' type scenarios. A downstream implication of limited data availability is that questions arise on the applicability of state-of-the-art methods to low-data volumes. Moreover, machine learning methods that are commonly used for low-data domains such as transfer learning [Pan and Yang, 2009] are not easily deployed to these novel domains due to the more rich sensor types that we require - such as multi-spectral imaging rather than standard Red Green Blue (RGB) imaging.

Given these challenges, in this paper we investigate the relative utility of applying more complex deep image processing techniques over what people might consider to be more baseline models for the specific domain of pastureland analysis. We make our investigation on the basis of a preliminary data collection that we performed to assess biomass and moisture content levels from grass prior to its use in producing silage. We proceed by first setting out some important elements from the state of the art regarding the visual assessment of crops and similar materials. We then outline the data collection methodology that we used to provide a basis for this initial study. Following this, we describe the image processing architecture that we apply, before detailing our results.

2 Non-Destructive Vegetation Monitoring

Estimating the quality of vegetation for silage production without expensive and destructive lab-based testing can be very beneficial. Fortunately, it has been found that the potential quality of silage can be influenced by specific factors including species mixture, moisture content, soil and crop condition, timing and weather, as well as the total biomass. From an imaging perspective, the texture and shape of a pasture canopy can help identify some of these characteristics [Abadi et al., 2010]. Images may show droplets of water on the leaves, or evidence of drought, seed heads or leaf size, while visual estimates of the canopy height is a strong indicator of biomass.

Classically the *visual*, or at least non-destructive, method for estimating the quality of vegetation has taken advantage of changes in electromagnetic reflectance values in vegetation that correlate with biophysical properties. Specifically, these approaches take advantage of the fact that as leaves photosynthesize they absorb most ultraviolet and visible light [Gobba, 2018] and around half of the available Near Infrared (NIR) light [Gitelson and Merzlyak, 1994]. The remainder is reflected. As the plant begins to senesce, or deteriorate with age, reflectance begins to decrease in the NIR and increase in the red bands. Different species and different conditions change the signature of reflected light [Gitelson et al., 1996]. Historically, researchers operationalised these facts into hand crafted functions, or Vegetation Indices (VIs), that are used as estimations of plant state. One of the most popular of these VIs is the Normalized Difference Vegetation Index (NDVI):

$$NDVI = \frac{R_{NIR} - R_{RED}}{R_{NIR} + R_{RED}} \quad (1)$$

where R_λ denotes reflectance at waveband λ [Xue and Su, 2017]. NDVI correlates to biomass early in the growing season, but saturates as vegetation becomes denser [Tanaka et al., 2015] and is influenced by soil exposure, topography, senescent vegetation and atmospheric contaminants. Canopy height has been used in conjunction with NDVI to extend its usefulness in dense vegetation [Schaefer and Lamb, 2016]. It is worth noting that NDVI like other VIs have been applied for many decades using remotely sensed data gathered from hand held devices through to data gathered from satellite mounted spectrometers [Townshend and Justice, 1988], as well as from high-resolution farm machinery mounted equipment [Schaefer and Lamb, 2016].

More recently, true image processing methods for vegetation analysis have started to become more popular in agriculture. One such automated system was developed to estimate vegetation cover and type using photos of 1m x 1m quadrats using Local Binary Patterns (LBPs) [McCool et al., 2018]. Over the past few years, the number of projects that use Deep Learning in agriculture has grown [Kamilaris and Prenafeta-Boldú, 2018]. These include estimating corn yields in the US using remotely sensed hyperspectral data and a deep neural

network [Kuwata and Shibasaki, 2016] and mapping winter vegetation quality coverage using Sentinel-1 SAR data and a recurring neural network [Minh et al., 2018]. Deep learning has also been used to estimate NDVI from Sentinel satellites, even on a cloudy day [Scarpa et al., 2018]. Ground-based images have also been used to estimate biomass at early growth stages. The performance of a Deep Convolutional Neural Network (DCNN) was compared to that of conventional models built on feature extraction such as image segmentation, colour comparison of R/G, B/G and 2G-R-B without segmenting images. The DCNN was more accurate and required no feature extraction prior to building the model [Ma et al., 2019].

3 Data Collection

To evaluate the feasibility of applying deep-learning methods to multi-spectral imaging of crops, we conducted a pilot study to gather data on grassland prior to silage production. Proximal visual and near-infrared (VIS-NIR) and LiDAR data samples were collected at three different venues in Ireland in Autumn 2018. We used a four-channel JAI AD-130 GE camera, mounted approximately 150cm above ground, pointing downwards. A LiDAR-Lite v3HP device was mounted beside the camera, pointing downwards, to record canopy height, while canopy height readings were also taken manually to ensure that the LiDAR height was measured correctly. Forty-six fully ground-truthed samples were collected. The data collected is summarized in Table 1.

Sensors and data collected		
Sensor	Purpose	Data Dimensions
JAI AD-130 GE	3-channel RGB image	964 x 1296 x 3
JAI AD-130 GE	1-channel NIR image	966 x 1296
LiDAR-Lite v3HP ²	Canopy height	Scalar (cm)
Weighing scales	Biomass	Scalar (Kg/Ha)
Oven	Dryness	Scalar (%age)

Table 1: Sensors and data format

In order to establish target values for biomass and moisture, manual measurements were taken on site, while sampling. This involved taking sensor readings of undisturbed grass at the target, applying a colour swatch and 50cm x 50cm quadrat on target and retaking sensor readings, before harvesting and bagging a 10cm x 10cm square within the quadrat. The bags were weighed to determine biomass. The contents were dried in an oven at 60°C for 24 hours and weighed again to determine the percentage of dry matter, henceforth known as dryness. Biomass ranged from 5144 to 60391Kg/Ha, with an average of 21129.22Kg/Ha and a standard deviation of 15532.58Kg/Ha. Dryness ranged from 14.3% to 43.6% with an average of 26.7% and a standard deviation of 9.6%.

Three samples were rejected due to full over-exposure of the NIR channel, resulting in a total of forty three acceptable samples. From the remainder, RGB and NIR images, LiDAR height, biomass and dryness estimates were available (see figures 1a and 1b for depictions of RGB and NIR data respectively). Each image was checked to ensure all components were present and then split into 48 patches of 156 x 156 pixels. The patches were checked for validity (no image that included unnatural features (such as trolley wheels or cables, quadrats, arms or legs) were considered to be valid). For a patch to be valid, both the RGB and NIR images had to be uncorrupted and not fully saturated. Each patch inherited LiDAR, biomass and dryness information from the full sample. Valid patches were combined (channels 1 to 3 are RGB, channel 4 is NIR) and flattened. These were saved along with the LiDAR reading, target biomass and dryness values for the sample to which they belong, to create a sub-sample. Once the sub-samples had been created, each was further checked by loading it and its labels, reverting to the original format and displaying them, as shown in Figure 1c. In order to avoid the situation where all sub-samples in a training batch were consecutive, therefore having the same ground truth, sub-sample order was randomized. Two datasets were prepared; a training dataset, with 1220 sub-samples and

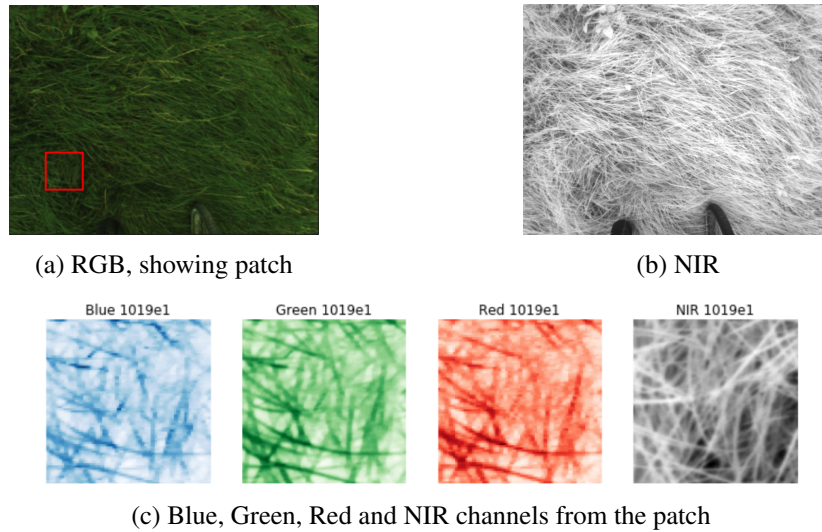


Figure 1: Sample images

a validation dataset with 306 sub-samples.

4 Data Modelling

Given two targets of biomass and dryness respectively, we developed a series of deep networks to estimate these scalars based on inputs from multi-spectral patch data and LiDAR height. In the following we briefly introduce the network investigated.

4.1 Baseline Convolutional Models

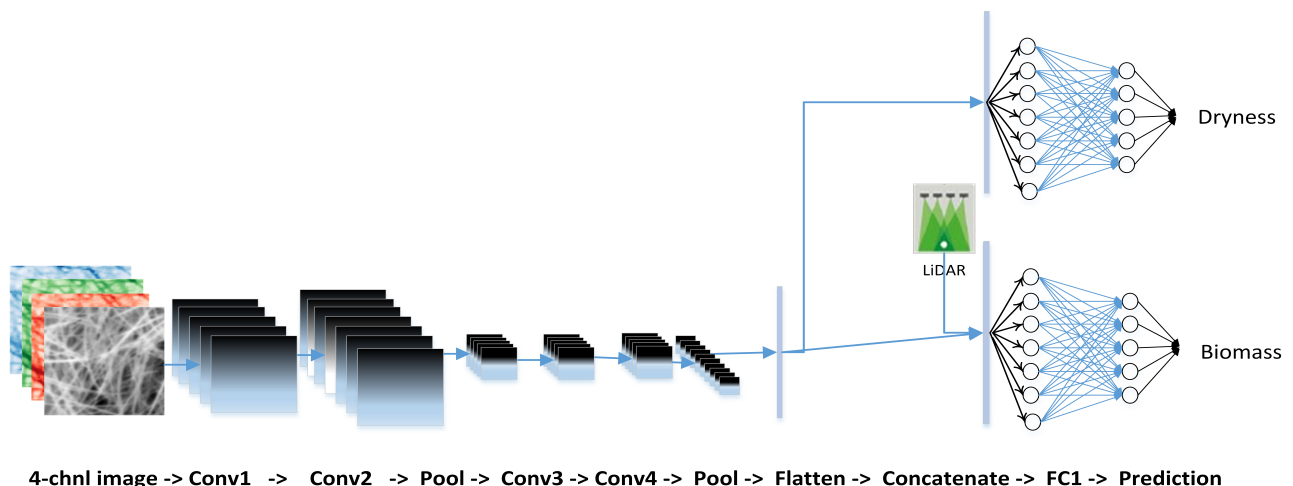


Figure 2: CNN 1 design

Our first three models are based around a convolutional neural network (CNN) architecture feeding into fully connected layers prior to a target layer as shown in Figure 2. The network depicted feeds four channel RGB and NIR patch data through four convolutional layers. Max-pooling was applied after the second and fourth convolutional layer respectively. The output of the second pooling layer was flattened and concatenated

with the scalar value from the LiDAR sensor before being fed through two fully connected layers. This CNN had inputs at two points and predicted two targets. These three baseline variants were:

- **CNN 1:** This is the network as depicted in figure 2. In the case of the target dryness, this network only uses the four-channel multi-spectral inputs. In the case of biomass, LiDAR data was concatenated to the outputs of convolution and pooling of the four-channel image data. This approach is based on the observation that LiDAR data has been used previously to enhance biomass estimates [Schaefer and Lamb, 2016], but has not been previously used to assist in estimation of dryness.
- **CNN 2:** This is a variant of CNN 1, which uses LiDAR data for both targets, so the output of the concatenation is fed into the first fully connected layer (FC1) in both cases. This variant gives the dryness target the option of learning from the LiDAR data, whilst the biomass target is trained as previously.
- **RGB CNN:** This CNN varies from CNN 1 and 2 in that rather than using four-channel image data, only the RGB channels were input to the convolution and pooling layers. As with CNN 1, the output was used exclusively to train for dryness, whilst LiDAR data was concatenated before training for biomass.

Our fourth and final CNN uses deep residual learning.

4.2 Deep Residual Network Modelling

While the CNN architectures outlined in section 4.1 provide a useful baseline, they arguably do not facilitate the true power of deep networks where highly complex functions of the input data can be learned automatically. Unfortunately as more convolutional layers are added to a network, the output becomes sparser and the ability to back-propagate error signal becomes more difficult, resulting in diminishing performance. To overcome this issue, residual networks use a skip-connection, where the output from an earlier layer is added to the output of a later layer, reinforcing the input to further layers. This reinforcement allows deeper networks to be built. Residual Networks reinforce learning by combining layers into blocks [He et al., 2016]. In this case, each block consists of a repeating stack of a convolution layer and a batch normalization layer, followed by an activation. The blocks are repeated, with an addition after each alternate block, just before the activation. Where the input changes shape, an extra convolution is added to ensure that the added input is the same shape. This technique is known as a skip-connection.

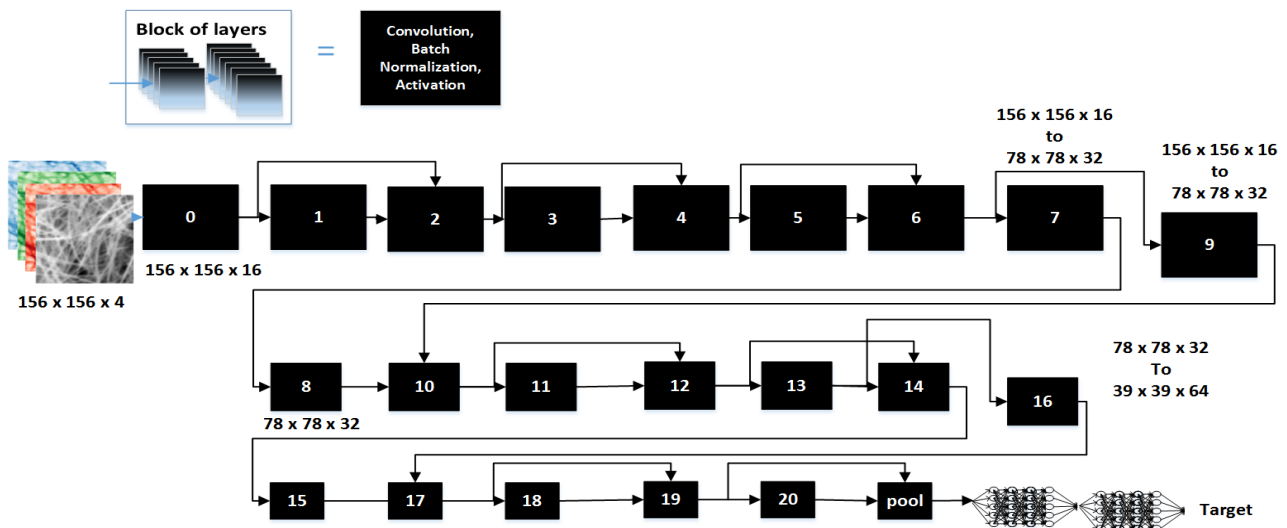


Figure 3: Residual Network Design

To leverage the full power of deep-learning to our pasture monitoring task, our fourth and final CNN uses a residual network to train a model to predict biomass and dryness. The network used here has a depth of 20 and

is depicted in Figure 3. In this case, two separate models were trained, one for biomass and one for dryness. No LiDAR data was used, but all four channels of the patch data were used.

4.3 Model Training

Given this is a regression problem, a mean square error loss function was applied. Models were trained for the number of epochs shown in table 2 and the validation dataset was evaluated against the model at the end of each training epoch. Models were implemented on a Keras front-end to Tensorflow and trained on Nvidia K40 GPUs.

5 Results

To evaluate the model we report the minimum Mean Absolute Percentage Error (MAPE) for both training and validation for both target variables and the four model variants outlined in the previous section. These results are summarized in Table 2. Focusing on the validation error in the baseline CNNs, we see that CNN 1 gave best performance for dryness, but performed slightly worse in the case of biomass. CNN 2 meanwhile performed extremely poorly for dryness but slightly outperformed CNN 1 for biomass. We see that by removing the near infrared information, that both biomass and dryness values were worse than is the case for CNN 1, underlying the usefulness of NIR data in this form of biological material analysis.

CNN	Epochs	Biomass		Dryness	
		Training	Validation	Training	Validation
CNN 1	1000	3.95%	45.34%	15.14%	22.15%
CNN 2	1000	3.02%	43.57%	23.88%	38.68%
RGB CNN	1000	4.18%	46.27%	18.8%	34.14%
ResNET	300	5.52%	17.96%	6.74%	9.86%

Table 2: Minimum Mean Absolute Percentage Error (MAPE) results

The most interesting results are clearly around the application of the residual networks. In both cases the validation error shows a significant improvement on all other model variants. While the validation error is of course not as low as the training error, it points to the usefulness of deeper models in vegetation analysis.

Given our scalar targets do not have a particularly wide range, it could be argued that the success of these results is due to networks converging towards a very small range of predictions for a given input. To validate the results with respect to this possibility we also calculate the precision error between actual values and a highly simplistic baseline prediction that is based on the averaged training value for a given target. Specifically, for each target (t), the average value from training data was calculated (\overline{T}_t). PE_t in this case was calculated as follows:

$$PE_t = \frac{1}{n} \sum_{i=1}^n \frac{|\overline{T}_t - V_t|}{\overline{T}_t} \quad (2)$$

where n is the number of test cases and V is the validation set. The results of this are shown in Table 3.

	Biomass	Dryness
Validation PE_t	52.28%	34.87%

Table 3: Training and validation data and results

These results for a simplified baseline demonstrate that even the basic CNN architectures are finding some useful signal in the data.

6 Conclusion

Although the number of samples used in developing this model is quite small, the results show a strong relationship can be established between the data gathered and both biomass and dryness. The use of a residual network showed very promising results, with the MAPE dropping substantially. Other factors that were considered were whether or not LiDAR was a useful inclusion for measuring moisture and whether the NIR channel is providing worthy extra information. While the data gathered and the analysis were centred on a pilot study, we believe the results show promise and validate the continued investigation of true deep learning methods in the analysis of data in an agri-food setting where data collection is expensive. Beyond expansion of the dataset, our future work will look at the relationship between patch size and reliability, as well as the reliability of results under different lighting conditions and with different model architectures.

Acknowledgments

The authors would like to thank those organizations that provided funding to enable this research. Patricia O’Byrne was supported by Enterprise Ireland Innovation Partnership IP 2018 0728 and Tanco Autowrap Ltd., and Robert Ross was supported by the ADAPT Centre for Digital Content Technology which is funded under the SFI Research Centres Programme (Grant 13/RC/2106) and is co-funded under the European Regional Development Fund. We would also like to thank Mark Deegan, Andre Rios, Mohammed Messabbah, Wael Rashwan and Fei Wang.

References

- [Abadi et al., 2010] Abadi, M., Capelle-Laizé, A.-S., Khoudeir, M., Combes, D., and Carré, S. (2010). Grassland Species Characterization for Plant Family Discrimination by Image Processing. In *Image and Signal Processing*, Lecture Notes in Computer Science, pages 173–181. Springer, Berlin, Heidelberg.
- [Gerland et al., 2014] Gerland, P., Raftery, A. E., Ševčíková, H., Li, N., Gu, D., Spoorenberg, T., Alkema, L., Fosdick, B. K., Chunn, J., Lalic, N., Bay, G., Buettner, T., Heilig, G. K., and Wilmoth, J. (2014). World Population Stabilization Unlikely This Century. *Science (New York, N.Y.)*, 346(6206):234–237.
- [Gitelson and Merzlyak, 1994] Gitelson, A. and Merzlyak, M. N. (1994). Quantitative estimation of chlorophyll-a using reflectance spectra: Experiments with autumn chestnut and maple leaves. *Journal of Photochemistry and Photobiology B: Biology*, 22(3):247–252.
- [Gitelson et al., 1996] Gitelson, A. A., Merzlyak, M. N., and Lichtenthaler, H. K. (1996). Detection of Red Edge Position and Chlorophyll Content by Reflectance Measurements Near 700 nm. *Journal of Plant Physiology*, 148(3):501–508.
- [Gobba, 2018] Gobba, F. (2018). 1616a Main factors influencing occupational solar uv exposure. *Occup Environ Med*, 75(Suppl 2):A86–A86.
- [He et al., 2016] He, K., Zhang, X., Ren, S., and Sun, J. (2016). Deep Residual Learning for Image Recognition. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016:770–778. arXiv: 1512.03385.
- [Kamilaris and Prenafeta-Boldú, 2018] Kamilaris, A. and Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and Electronics in Agriculture*, 147:70–90.
- [Kummu et al., 2016] Kummu, M., Guillaume, J. H. A., de Moel, H., Eisner, S., Flörke, M., Porkka, M., Siebert, S., Veldkamp, T. I., and Ward, P. J. (2016). The world’s road to water scarcity: shortage and stress in the 20th century and pathways towards sustainability. *Scientific reports*, 6:38495.

- [Kung Jr, 2010] Kung Jr, L. (2010). Understanding the biology of silage preservation to maximize quality and protect the environment. In *Proceedings, 2010 California Alfalfa & Forage Symposium and Corn/Cereal Silage Conference*, pages 1–2, Visalia, CA.
- [Kuwata and Shibasaki, 2016] Kuwata, K. and Shibasaki, R. (2016). Estimating Corn Yield in the United States with MODIS EVI and Machine Learning Methods. In *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, volume III-8, pages 131–136. Copernicus GmbH.
- [Kyte and Kaldani, 2013] Kyte, R. and Kaldani, R. (2013). World Bank Group Agriculture Action Plan 2013-2015. Technical report, The World Bank Group, Washington DC. [Online] Available at: <http://www-wds.worldbank.org/> (accessed 2016).
- [Ma et al., 2019] Ma, J., Li, Y., Chen, Y., Du, K., Zheng, F., and Zhang, L. (2019). Estimating above ground biomass of winter wheat at early growth stages using digital images and deep convolutional neural network. *European Journal of Agronomy*, 103(2019):117–129.
- [McCool et al., 2018] McCool, C., Beattie, J., Milford, M., Bakker, J. D., Moore, J. L., and Firn, J. (2018). Automating analysis of vegetation with computer vision: Cover estimates and classification. *Ecology and Evolution*, 8(12):6005–6015.
- [Minh et al., 2018] Minh, D. H. T., Ienco, D., Gaetano, R., Lalande, N., Ndikumana, E., Osman, F., and Maurel, P. (2018). Deep Recurrent Neural Networks for Winter Vegetation Quality Mapping via Multitemporal SAR Sentinel-1. *IEEE Geoscience and Remote Sensing Letters*, 15(3):464–468.
- [Pan and Yang, 2009] Pan, S. J. and Yang, Q. (2009). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359.
- [Scarpa et al., 2018] Scarpa, G., Gargiulo, M., Mazza, A., and Gaetano, R. (2018). A CNN-Based Fusion Method for Feature Extraction from Sentinel Data. *Remote Sensing*, 10(2):236.
- [Schaefer and Lamb, 2016] Schaefer, M. T. and Lamb, D. W. (2016). A Combination of Plant NDVI and LiDAR Measurements Improve the Estimation of Pasture Biomass in Tall Fescue (*Festuca arundinacea* var. Fletcher). *Remote Sensing*, 8(2):109.
- [Tan et al., 2014] Tan, A. E., Richards, S., Sarrabezolles, L., Platt, I., and Woodhead, I. (2014). Proximal soil moisture sensing of dairy pasture. In *2014 IEEE Conference on Antenna Measurements Applications (CAMA)*, pages 1–4.
- [Tanaka et al., 2015] Tanaka, S., Kawamura, K., Maki, M., Muramoto, Y., Yoshida, K., and Akiyama, T. (2015). Spectral Index for Quantifying Leaf Area Index of Winter Wheat by Field Hyperspectral Measurements: A Case Study in Gifu Prefecture, Central Japan. *Remote Sensing*, 7(5):5329–5346.
- [Townshend and Justice, 1988] Townshend, J. R. G. and Justice, C. O. (1988). Selecting the spatial resolution of satellite sensors required for global monitoring of land transformations. *International Journal of Remote Sensing*, 9(2):187–236.
- [UN, 2015] UN (2015). World Population Prospects The 2015 Revision Key Findings and advance tables. Technical Report ESA/P/WP.241, United Nations, Department of Economic and Social Affairs, Population Division, New York City. United Nations, Department of Economic and Social Affairs, Population Division [Online] Available at http://esa.un.org/unpd/wpp/publications/files/key_findings_wpp_2015.pdf (Accessed 2016).
- [Xue and Su, 2017] Xue, J. and Su, B. (2017). Significant Remote Sensing Vegetation Indices: A Review of Developments and Applications. *Journal of Sensors*.