USING MACHINE LEARNING TO IDENTIFY DIFFERENT SPECIES OF TIMBER, USING IMAGES OF FACE GRAIN AND END GRAIN TIMBER, AND COMPARING THE RESULTS.

Sean Hennessy: C00138397
Supervisor: Dr Oisin Cawley
M.Sc. in Data Science

1. Introduction

- Identifying different timber species can prove challenging sometimes. There are different methods for classifying different timber species. According to (Desch and Dinwoodie, 1996), traditionally these have included:
  - Scientific analysis, specifically DNA testing, chemical testing and spectroscopy.
  - Inspection of timber properties, such as; colour, smell and grain.
  - Visual examination of the end-grain structure macroscopically (10X magnifying lens) and microscopically, by wood anatomists.
- Illegal timber logging and trade in some countries has led to the need for rapid identification of timber species in the field by government agencies.
- The scientific techniques are not suitable for the field and there is a shortage of wood anatomists.
- This has led to machine learning algorithms being used to classify timber species.
- There's a free mobile application for Malaysian timber species.
- There is also a commercial product aimed at government agencies, which can identify a lot of North American and European timber species. However it is too expensive for regular woodworkers.
- It would be useful to have a free or low-cost mobile or computer-based application which could classify some of these timber species.
- The development of such an application is beyond the scope of this project.
- Instead this project is intended as a proof of concept, to see if a Machine Learning algorithm, specifically a Convolutional Neural Network (CNN) can be used to successfully classify three different species of timber from photographs.

2. Research Questions

- Can a CNN be retrained using transfer learning to successfully classify three different species of timber from photographs?
- Will the CNN classify photographs of these types of timber with a greater degree of accuracy using face grain or end grain photographs?
- The accuracy of the retrained CNN will be compared with the literature.

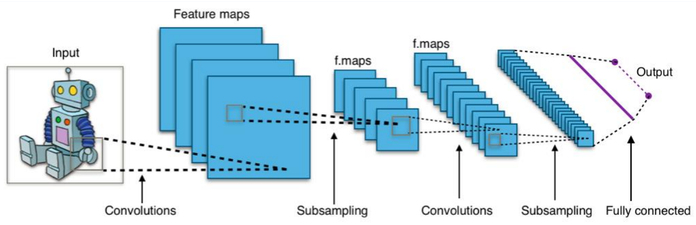


Figure 1: Diagram showing how CNNs work (towards data science, 2019).

3. Literature Review

- The use of artificial intelligence (AI) has increased greatly in the last few years, in academic research and industry.
- The recent advances in AI are mostly due to deep learning (DL) which is a subfield of machine learning (ML), which is a subfield of AI.
- Traditionally ML involved pre-processing data to create useful features which were then fed into ML algorithms, like; random forests (RF) or support vector machines (SVMs). This required domain knowledge to do successfully.
- CNNs are a popular DL technique, which involve networks of artificial neurons in multiple layers, which can detect features automatically, thus avoiding having to manually create features, and they can be used for classifying huge amounts of data.
- Also CNNs trained on one dataset, can be applied to other datasets, using transfer learning and finetuning or training the top layers of the CNN.
- Previous research has been carried out on timber classification.
- (Vácha and Haindl, 2013) analysed timber images using Markevian textural features to match them with database images.
- (Tang et al., 2018) used a CNN called SqueezeNet to identify 100 species of Malaysian timber, and help tackle illegal logging and trade, and produced an iOS mobile application – MyWoodPremium.
- XyloTron is a product developed to help combat illegal logging and trade. It processes timber images through wavelet analysis and partial least squares methods to create a morphological signature, which can be compared to reference images to find the best match (Xylotron.org, 2019).

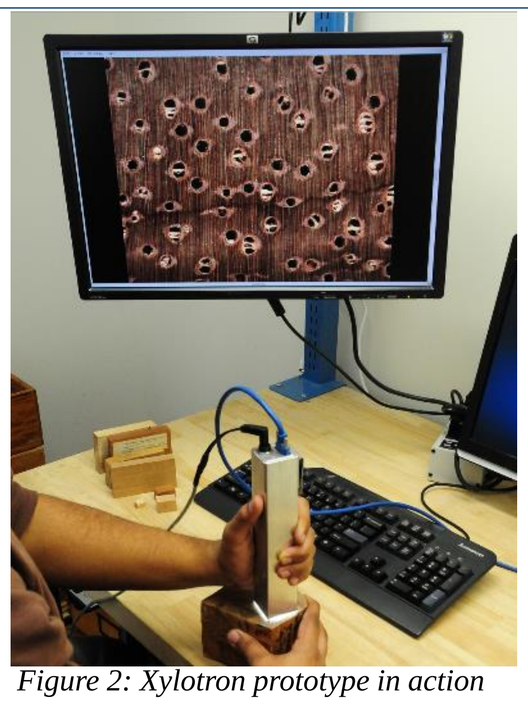
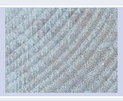


Figure 2: Xylotron prototype in action (USDA, n.d.).

4. Datasets

- The datasets were produced by taking face grain and end grain photographs of Irish Common Ash (*Fraxinus excelsior*), American Sugar Maple (*Acer saccharum*), and Norway Spruce (*Picea abies*).

Table 1: Details of the Face Grain and End Grain Datasets.

Dataset	Timber Species	Number	Image
Face grain	Ash	634	
Face grain	American Sugar Maple	611	
Face Grain	Norway Spruce	727	
Face Grain	All Species	Total = 1972	
End Grain	Ash	618	
End Grain	American Sugar Maple	653	
End Grain	Norway Spruce	607	
End Grain	All Species	Total = 1878	

5. Research Methodology

- Face grain and end grain photographs were taken of the Irish Common Ash, American Sugar Maple, and Norway Spruce.
- The face grain dataset and end grain dataset will be split into a training set, test set and validation set.
- A CNN will be trained using transfer learning to classify the three timber species.
- The accuracy of the classification using face grain photographs will be compared to the accuracy using end grain photographs.
- Possible technologies that could be used include; python, a CNN like SqueezeNet, TensorFlow, Keras, R, the IT Carlow Hadoop cluster and Apache Zeppelin notebooks.

6. Next Steps

- If the CNN successfully classified the timber species with a high level of accuracy, it could be trained to classify more timber species.
- It can be established whether it is better to use face grain or end grain photographs for classification.
- A free or low-cost mobile or computer-based application could be developed and made available for regular woodworkers to help in identifying timber species.

References

Desch, H. and Dinwoodie, J. (1996). Identification of Timbers. In: H. Desch and J. Dinwoodie, ed., *Timber Structure, Properties, Conversion and Use*, 7th ed. London: Palgrave, pp.56-65.
towards data science, 2019. *Diagram Showing How CNN's Work*. [image] Available at: <https://towardsdatascience.com/introducing-convolutional-neural-networks-in-deep-learning-400f9c3ad5e9> [Accessed 5 April 2020].
USDA, n.d. *The Xylotron Prototype In Action*. [image] Available at: <http://tefso.org/download/Other-en/ENG-XyloTron-introduction.pdf> [Accessed 5 April 2020].
Vácha, P. and Haindl, M. (2013). Wood Variety Recognition on Mobile Devices. *ERCIM News*. [online] Available at: https://ercim-news.ercim.eu/en93/ri/wood-variety-recognition-on-mobile-devices [Accessed 5 April 2020].
Tang, X., Tay, Y., Siam, N. and Lim, S. (2018). MyWood-ID: Automated Macroscopic Wood Identification System using Smartphone and macro-lens. In: *CIIS 2018: Proceedings of the 2018 International Conference on Computational Intelligence and Intelligent Systems*. [online] New York: Association for Computing Machinery, pp.37–43. Available at: https://dl.acm.org/doi/10.1145/3293475.3293493 [Accessed 5 April 2020].
Vácha, P.
Xylotron.org. (2019). Xylotron, wood identification with machine vision technology. [online] Available at: https://www.xylotron.org/XT-Technology.html [Accessed 5 April 2020].