AUTOMATING ARCHITECTURAL DRAWING CLASSIFICATION:

Implementing a Convolutional Neural Network (CNN) to classify between Plans and Sections

M.Y. NG,
University of New South Wales, Sydney, Australia
jennmyng@gmail.com

Abstract. (X) There is an accumulation of digital architectural drawings as the industry has progressively entered the digital realm, moreover, the files are often unstructured and uncategorised resulting a struggle in file searching, leading to an inefficient workplace. To resolve the issue, this research project intends to take a machine learning approach, aiming to automate a image classification process through a trained system - Convolutional Neural Network (CNN), constructed in Python. CNN is a type of artificial neural network, branching from machine learning. It is known for its spacial recognition, best for classifying images. It analyses every pixel to output image patterns and features allowing it categorise them apart. The neural network is trained and validated through different iterations through tweaking hyperparameters with hundreds images of plans and sections drawings. The outcome of this project aims to provides a well-functioning CNN with an classification accuracy rated above 90% . In future works, many companies can start to implement section details and scanned old hand-drawn files to be included in the system with ease, to enrich the resource library for referencing. This research project is aimed to encourage a higher efficiency within design workplaces and an improved file archiving system. Furthermore, the model can be further trained to identify more specific features such as abstract drawings, texts and other visualisations; eventually be complex enough to apply on other documents. Ultimately, the findings of this research project could expand the possibility to be adopted by other disciplines and practices.

Keywords. Convolutional Neural Network, machine learning, image classification, architectural drawings, filing

1. Introduction

Artificial Intelligence (AI), through popular accounts, is often portrayed as futuristic and is focused on dystopic outcomes and its threat to society. (Bostrom, 2014) However, machine learning (ML), a field of AI that is highly integrated into our daily lives (Ayers, 2016). From the micro-scale such as an individual's smart phones, home and security systems and personalised social media feeds, to a macro scale such as transportation and cyborg technology. Moreover, ML is available to help lessening repetitive manual tasks which also requires sophisticated understanding on the subject, can transfer individual's focus onto more complex task and increase work efficiency.

Within the built environment discipline, ML can and is not limited to be implemented within design processes but to everything that revolves around design. For example, workplaces' dynamics in an office (Phelan, 2016) can be maximised to guide designers understanding their working space thoroughly, proceed to a more efficient environment. In the project, room usage data were collected and learnt through ML to better predict room usage, ultimately giving its staff the most efficicent room for an improved working experience.

For this research project, it is in collaboration with Arup, an international Engineering and Design firm, together we are motivated to tackle a real-life problem, similarly aiming to create a more efficient workplace under the aid of ML. As architectural drawings have entered the digital realm it results an accumulation of files. However, these files are often unstructured and unsystematic, subsequently Arup's engineers are manually executing file sorting themselves repetitively. With ML coming into place, it can off-load engineers doing repetitive tasks and undertake a sophisticated visual understanding to distinguish targeted images.

2. Research Aim

Ultimately, with the research project, it is aimed to prove the potential of machine learning, through constructed model that can successfully predict different kinds of architectural drawings such as plans and sections. To achieve the aim, there are two objectives that needs to be met.

First, understanding the concept of ML and finding out which type is the best to solve this problem is essential in order having the most efficient outcome.

Second, a model is aimed to be constructed that can effectively help predict given images, such as plans and sections, under optimal hyperparameters that will be trained and validated through multiple iterations.

3. Research Questions

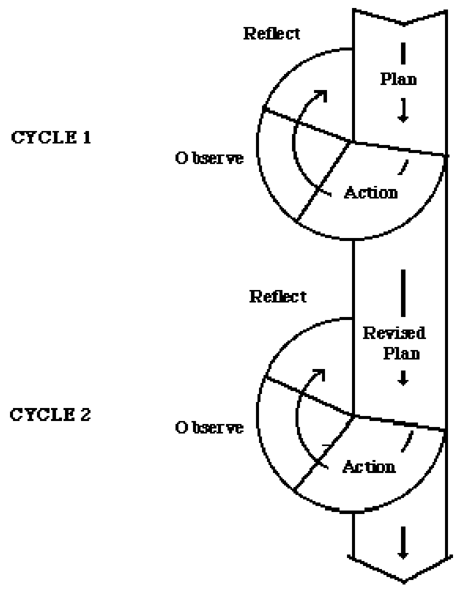
Ultimately, the research project revolves around a main question: "How can drawing classification be automated within the built environment discipline using machine learning?"

In order to have sufficient research to answer, questions in regards to concepts, aim and methodology for clarification can aid lead to the main question. For example, "What is the difference between ML and Neural Networks?", "How are training, validation and prediction different from one another, and how can they provide the correct data to help improving the next iteration?"

4. Methodology

In carrying out this research project, an action research approach is applied.

Action research, stated by D. Gabel (1995), it involves a characteristic cycle: Exploratory Stance, Action, Observation and the cycle continues until the aim is eventually achieved.



[FIGURE 1]

Figure 1. Diagram of Action Research (Hopkins, 1985)

In the Exploratory Stance stage, the impact of AI and ML on daily lives and design processes have been discussed. With a real-life problem addressed by Arup, highlighting a need of ML in solving file sorting in the built environment discipline. This has provided an opportunity to dive deeper in exploring.

Stage two, Action, is “carrying out the intervention” (D. Gabel, 1995), this includes learning ML fundamentals, understanding the concepts and grasping what is important to be implemented into this specific project.

Stage three, Observation, it involves a big role of communication where “everyone has empowerment” (D. Gabel, 1995). Not only the project is researched and tested alone, other parties such as industry partners, tutors and peers are involved in the process as the observants for contribution. With the presence of industry partner, they provide the deepest insight to the problem and offer feedback to improve the approach of the project.

Lastly, the cycle of completing all the steps are done until all desired is achieved. In the project, many testing and validations on different iterations will be done to test the robustness of the solution. Can it be adapted to different architects’ style of plans and sections, can it be adapted to different line thicknesses, colours and texts, can this project be further expanded in future studies for students in similar disciplines; or different? All these questions are always reflected on during evaluation, feedback done by tutors and peers, ultimately practically tested through technical skills.

5. Background Research

5.1. EVOLUTIONAL CHANGE IN ARCHITECTURAL DESIGN PROCESS

The design process in architecture has evolved and continues to evolve to this day with the fast progression of technological advancements. From manual drawing tools such as a pencil to digital tools such as Grasshopper, a visualized scripting platform “… the discipline had yet to complete the portrayed revolutionary cycle…” (Wirz, 2014) However, thanks to researchers and architects who have engaged in the tools and shared their experience through open-source platforms such as online tutorials and forums, the new computational tools are more accessible to individuals especially students. “This gives students, coming from a more conservative background, the opportunity to practically understand this new approach of design.” (Wirz, 2014) As a result, the new emergence of computational tools is becoming more accessible and ultimately can expand the potential in designing in architecture.

Peter Eisenman’s “Biozentrum” (1987), a biology centre for Goethe University in Frankfurt is a great example of how digital drawing tools has expanded boundaries within architectural design processes. The fundamental concept of the design references DNA sequences and Eisenman "… use(d) a computer as a procedural modelling tool capable of drafting predefined figures at varying alignments…" (Zardini, 2013) This highlights (X)

Moreover, the rise of variations in computational methods has proved more than purely focusing on the aesthetic aspects in architecture, but also the functionality and technicality of the projects. Wirz continues to emphasise, "Digital design can and must confront reality as early as in concept stages giving answers to both performance and functional criteria…" (Wirz, 2014) He strongly agrees the need of cross-disciplinary culture in the discipline and see it as a way of expanding the possibilities to achieve more, "form must reflect the integration between multiple disciplines." Thus, not only the tools are more accessible, but it also highlights the importance of engaging greater knowledge across different disciplines to optimise design.

For this research project, the article significantly explores and verifies the potential of computational tools used within the architectural discipline. Not only it helps to refine and enhance the design process for the designer, but can it do more? Ultimately his perspective encourages this project to experiment outside the box and hope to future optimise design processes within architecture.

5.2 FITTING ML INTO THE CHANGED PARADIGM

Artificial Neural Networks (ANNs) has been introduced in the last ten years. It sits under the machine learning paradigm but referenced from a biological neuron system. The network can mimic information processing and acquire knowledge like a human brain, "ANNs are more robust and rapid than traditional techniques", wrote I.A. Basheer and M.Hajmeer in their paper "Artificial Neural Networks: Fundamentals, Computing, Design, and Application". The article explains how the network learns by familiarising the readers with the structures of neurons within the human brain. Furthermore, in the architectural design processes, the idea of automation and machine learning within the discipline is encouraged, wish to expand design possibilities, within and outside the design process.

A paper published by Leon Gatys, Alexander Ecker and Matthias Bethge, "A Neural Algorithm of Artistic Style", the authors explores the potentiality of Convolutional Neural Networks (CNNs), a type of network that is commonly used in processing imagery such as image recognition and image-style transfer. The authors' research explains the network dismantling the image in hierarchal layers for recognising pixels and lastly identifying certain parts to conduct style transfer. Aligning this idea to the potential in architectural design process, the idea of image recognition and style transfer can potentially help in recognising plans, sections and elevations in help to automate processes for architects.

5.3 CONVOLUTIONAL NEURAL NETWORK (CNN)

ML revolves around the way it learns by themselves, involving self-changing algorithms within the learning process of an input set of data. Within ML, neural networks is a series of algorithms to also train computers, but to mimic the system of a human brain – to recognise pattern and group input data into categories. Eventually, as the algorithm further sophisticates, it is known as "deep neural networks".

For this research project, it is specific to research a solution to be able to recognize imagery and spatial patterns from architectural drawings, such as plans and sections. Before constructing a neural network model, it is important to understand the fundamental structure. In 1980, K. Fukushima has published a paper "Neocognitron: A Self-organizing Neural Network Model for a Machanism of Pattern Recognition Unaffected by Shift in Position", about improving a neural network model's ability to recognize patterns, nicknamed "neocognitron"; an extended self-organising multilayered neural network "cognitron"..

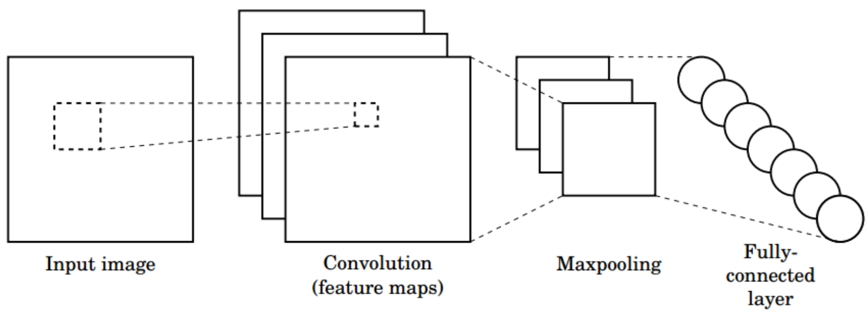
The structure of the network proposed in the paper is similar to the proposed hierarchy model of the visual nervous system by Hubel and Wiesel (1962, 1965). It is structured flowing from simple cells to complex cells. With Fukushima's proposal, the to-be-improved network is aimed to handle position variations within input data. Opposed to previous versions of networks, they are sensitive to the position and found difficulties in normalising it.

To improve a position-invariant pattern recognition, the structure of the network consists of C and S cell layers, each sub-grouped into "cell planes", and each cells have receptive fields of the same function of different positions. Most importantly, with the layers receiving the same set of input synapses, parameter alternations are the determinants for the network to differentiate between similarities and distortion of the patterns.

Within the computer simulation done by Fukushima, despite the input shared similar features amongst each other, such as "X" and "Y" sharing identical upper parts plus "X" and "Z" sharing the same diagonal inclination, the network has "acquired the ability to discriminate them correctly" (K. Fukumisha, 1980) due to repetitive feeding of the resembling patterns. Ultimately, it is concluded that the increase of cell planes would "presume" to increase recognition steadiness and accuracies. Due to Fukushima's computer's lack of memory, the layers could not be increased. With the paper's prove of concept that a neural network can achieve position-invariant recognition and now a higher memory compacity computer is accessible, this project can be taken further.

6. Case Study (1000-1500w)

It is understood that CNN is known for image recognition. Compared to other neural networks, it is capable of spatial awareness, meaning the network recognises the relationship between each pixel and is able to generate patterns. From Figure 1, it is a simplified diagram showcasing the structure of CNN.



[FIGURE 2]

Figure 2. Diagram of Convolutional Neural Network (Pavlovsky, 2017)

6.1 INPUT LAYER

Images of the same category will be input into the model for training and validation. Each image will be broken down into matrix of input values dependent on their RBG values. For this model, 800 images will be input for training and 200 images will be input for validation.

6.2 CONVOLUTIONAL LAYERS

In Convolutional Layers, each image pixels are analysed by filters called *kernels.* Every time a filter goes pass each images by *n* strife value, it contains a grid of randomised values overlaying on top of the image pixels. The overlapped numbers will be multiplied, resulting an output. (Kernel size and quantities are set by the designer.)

6.3 MAXPOOLING

The next stage after collecting outputs from the feature maps. Maxpooling helps to reduce image size and keeping the highest value at each kernel stop. Ultimately, this aims to reduce training time and prevent overfitting by cancelling out small details.

6.4 FULLY CONNECTED LAYERS

Last stage of the model; these layers flatten the feature maps that has been generated previously, while connecting every value to train, validate and predict the output data.

6.5 TRAINING & VALIDATION

Input images will go through the structure explained above for training and validation. Hyperparameters are constantly changed for iterations in order to find the best version of the model, that is robust enough for the project.

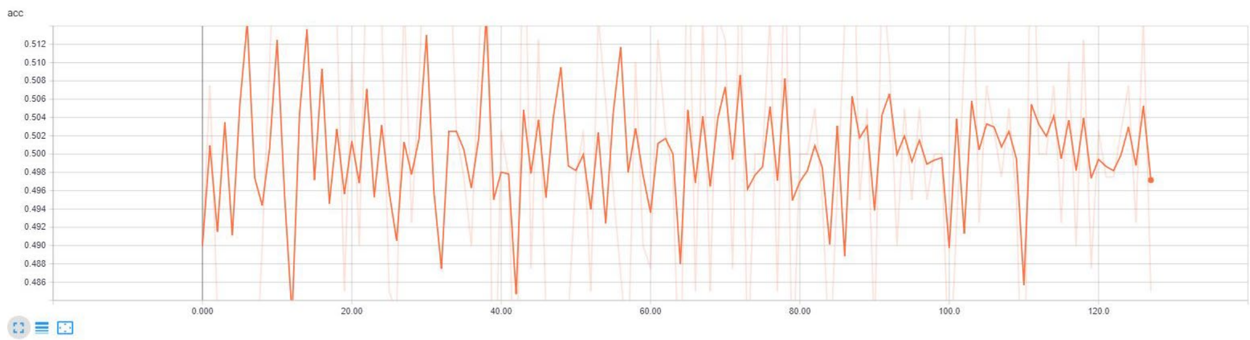
6.5.1 Hyperparameters

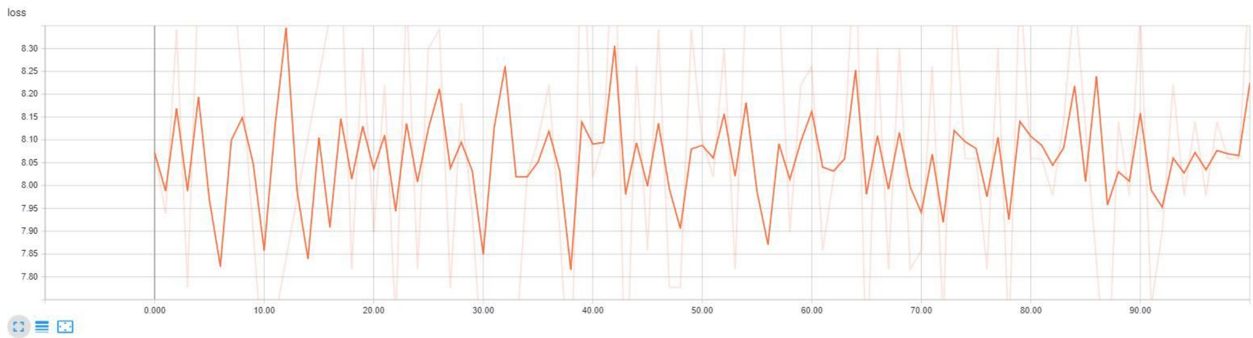
Hyperparameters to change	**Purpose**
Convolutional Layer numbers	How detailed the model should capture the given images
Number of filters	To determine how complex the model should analyse the input data
Dropout Percentage	To train the model’s dependency on its training data
Learning Rate	To adjust how quickly the model can arrive at the best accuracy
Epoch	To determine how many times the input data go through the model

6.5.2 Changing Hyperparameters

During the process of finding the model, alteration of the hyperparameters is done to achieve the optimal model for this project. A few iterations were created, and three are selected to be evaluated in detail in this paper. During evaluation of each iterations, it is important to look for a high percentage of accuracy and a low percentage of loss. Dips and fluctuation in the graphs could mean overfitting, meaning the model is corresponding too closely to every single data point, resulting a bad formation of the pattern trend.

1. Convolutional Layer numbers: 1
   Number of filters: 32
   Dropout Percentage: 0.2
   Learning Rate: 0.001
   Epoch: 128



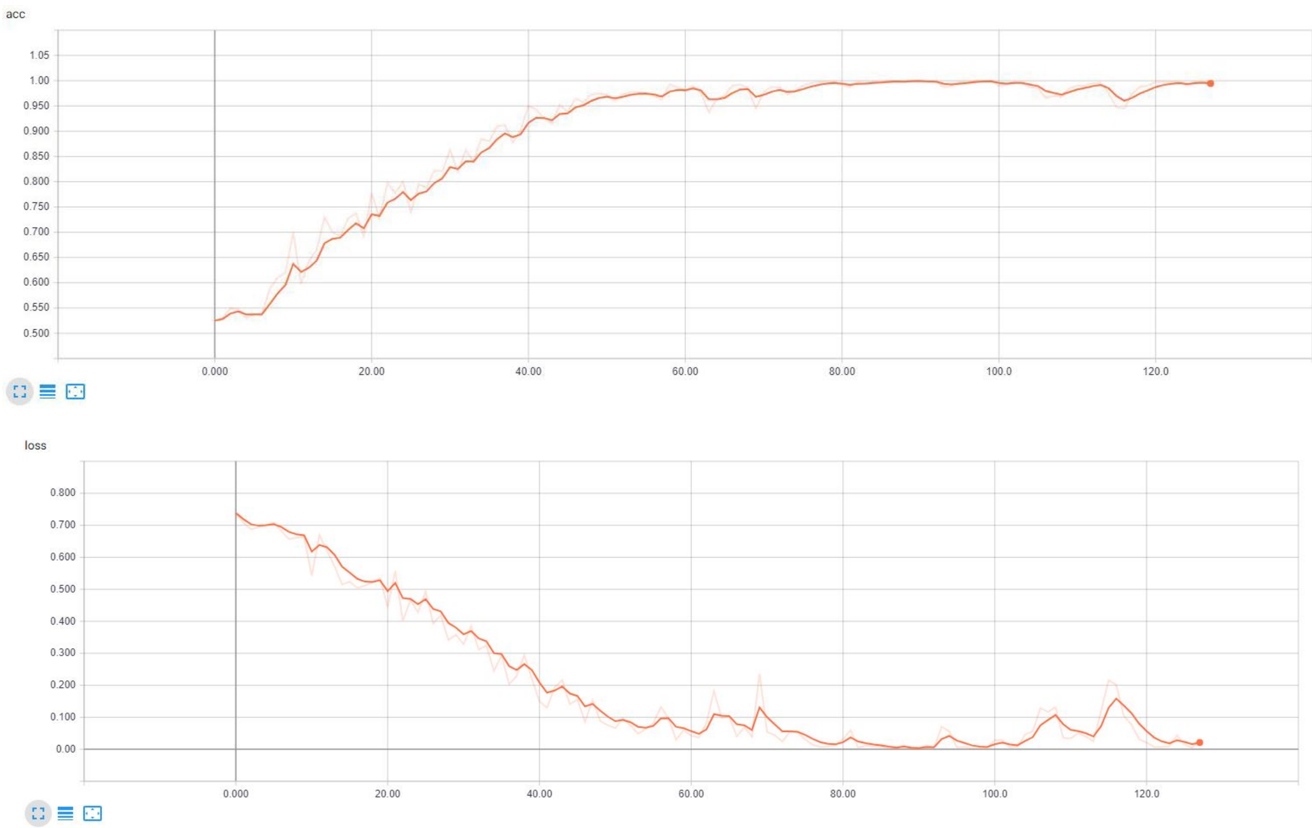


[FIGURE 3, 4]

Figure 3, 4. Training Accuracy and Loss percentage graphs of Iteration 1

During evaluation of the two graphs above, the accuracy graph has reached a top of 52%. However, this is not ideal as an optimal figure to prove a well-trained model is at least above 90%. Similarly with the loss graph, 0% is the aim therefore a minima of 7.2% is not considered ideal. The reason of a poor result could be the model not picking up enough features from the input images therefore unable to categories them effectively. Therefore, to improve the results, 2 more layers are added to the convolutional layers in Iteration 2.

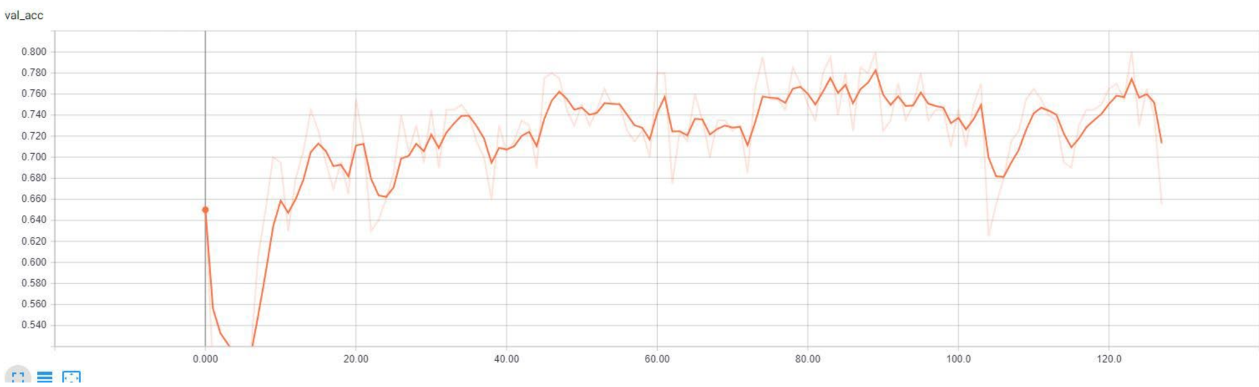
2. Convolutional Layer numbers: **3**
   Number of filters: 32
   Dropout Percentage: 0.2
   Learning Rate: 0.001
   Epoch: 128

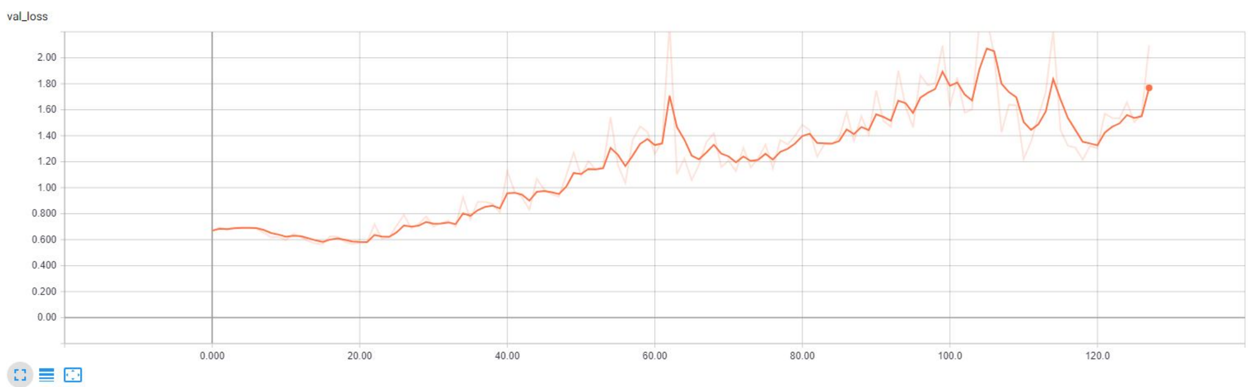


[FIGURE 5, 6]

Figure 5, 6. Training Accuracy and Loss percentage graphs of Iteration 2

In Iteration 2, the training accuracies and losses results are ideal. However, in comparison to its validation results:



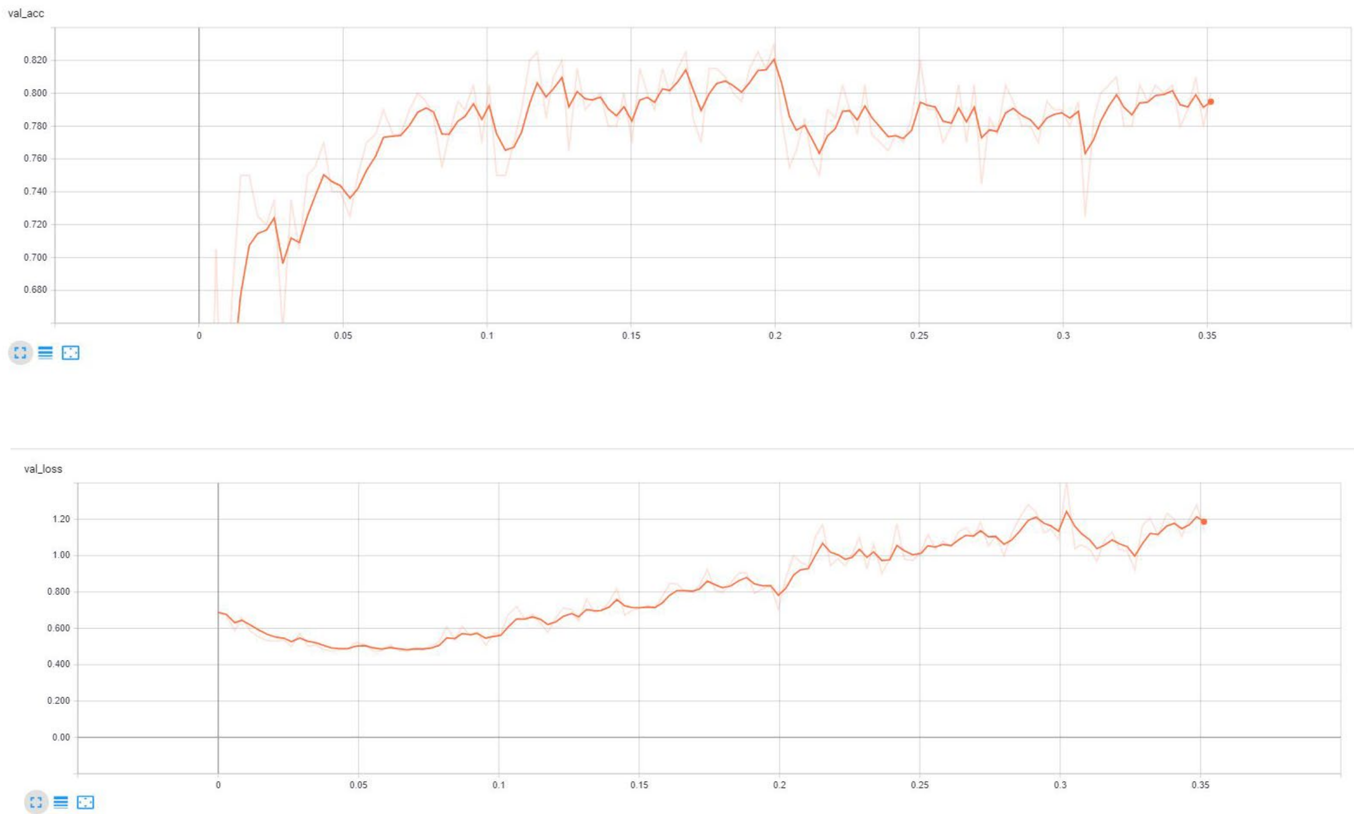


[FIGURE 7, 8]

Figure 5, 6. Validation Accuracy and Loss percentage graphs of Iteration 2

The validation results of Iteration 2 appear to have a lot of dips and fluctuation. These could mean a high possibility of overfitting and the model would not give a reliable prediction due to the inability to generate the correct pattern trend. Therefore in Iteration 3, the learning rate of the model is lowered for improvement.

3. Convolutional Layer numbers: **3**
   Number of filters: 32
   Dropout Percentage: 0.2
   Learning Rate: **0.0001**
   Epoch: 128



[FIGURE 9, 10]

Figure 9, 10. Validation Accuracy and Loss percentage graphs of Iteration 3

Reading the validation result graphs from Iteration 3, although there are still fluctuations, they are more stable. The validation accuracy reached a high of 82% and validation loss managed to hit a low of 0.5%. With the limitation of time, model from Iteration 3 is taken for prediction.

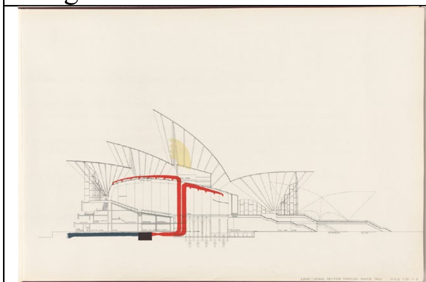
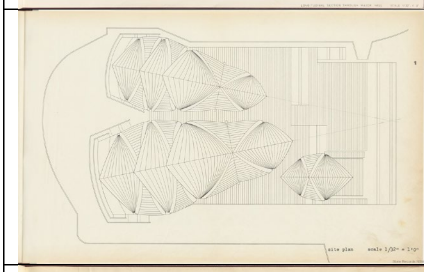
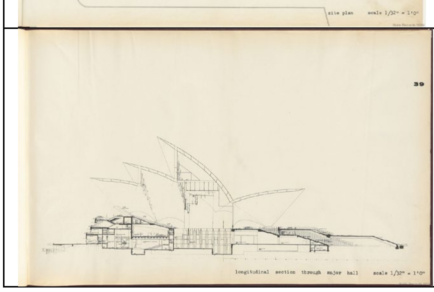
6.3 PREDICTION

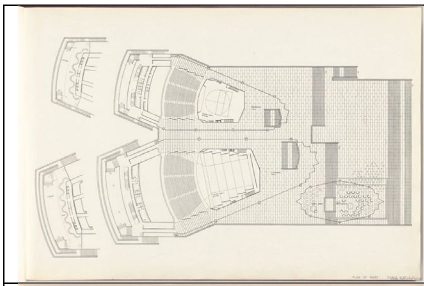
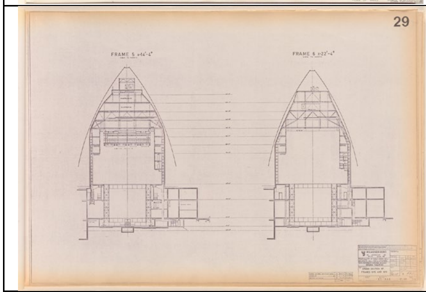
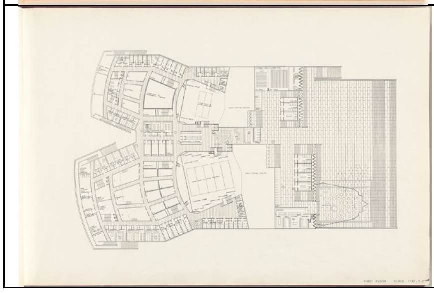
New images are used as input images to test the robustness of the model. Opera house plans and sections were chosen for the architecture's curvature and unique design that are not often seen in majority of drawings.

```
Image: 12707_00050.jpg... Prediction: Section architecture with 100.00% confidence.
Image: 12708_00003.jpg... Prediction: Plan architecture with 93.48% confidence.
Image: 12708_00041.jpg... Prediction: Section architecture with 100.00% confidence.
Image: First_Floor_Plan_(Sydney_Opera_House)_(5373921522).jpg... Prediction: Plan architecture with 99.72% confidence.
Image: NRS12800-2_a122_000029.jpg... Prediction: Section architecture with 99.85% confidence.
Image: Sydney-Opera_planta_vig-1024x695.jpg... Prediction: Plan architecture with 99.58% confidence.
```

[FIGURE 11]

Figure 11. Prediction results using model of Iteration 3

Image	Result
	Section, confidence 100%
	Plan, confidence 93.48%
	Section, confidence 100%

	Plan, confidence 99.72%
	Section, confidence 99.85%
	Plan, confidence 99.58%

The model has successfully predicted all the images correctly. However, looking at the high percentages, it can mean it is heavily reliant on the training data and undermining a chance of overfitting. Meaning further finetuning is needed before official implementation.

7. Significance of Research

The set aim of this project is to prove the potential of machine learning, through constructed model that can successfully predict different kinds of architectural drawings such as plans and sections. Through the research project, the model has successfully predicted results over 90% confidence, proving the possibility to take on further in development.

Not only Arup encounters an abundance of unstructured and uncategorised digital drawing files, the engineers do often reference back to old hand-drawn drawings that are still in physical files. In wish to meet the full digitalised future, the physical files are aimed to be scanned, recognised and be able to searched by engineers easily.

Therefore with the help of the developed model, it holds a great significance in setting a prove of concept to the possibility of machine learning being implemented to help within the design workplace, improving efficiency; rather being “a threat to humanity”.

Moreover, through the process of constructing the model, its hyperparameters were constantly tweaks to provide the best result possible. This is a significance finding as it highlights an important feature of ML: each ML model is project-specific. This definitely breaks down the generalisation of ML in the society of it being a whole system on its own.

8. Evaluation of research project

This research project aims to automate architectural drawing classification, by handling a wide range of design styles. This project is partnered with Arup and solve an existing problem, where company hopes to find a way to classify plans and sections automatically. In the process, learning about CNN (Convolutional Neural Network) and implementing it into Python has allowed a solid foundation to develop the model effectively, and allow a deep understanding to help debug within the code. Training the model requires tweaking hyperparameters within the code such as changing layers, dropout rates and learning rates. Although “ANNs are more robust and rapid than traditional techniques”, wrote Basheer, through multiple iterations of trained models, it reveals that training a neural network to an optimal prediction rate, is very project specific and the hyperparameters cannot be generalised despite being assumed by majorities. The model can classify and provide a prediction percentage of the input images, it successfully outputs an above 90% prediction from simple and sophisticated drawings. The project goes in depth with the possibility of a neural network and what it can achieve, such as learning how to classify images despite variations in size, line thickness and styles of design. Due to time limitation, the model cannot be implemented into the company’s system. However, the concept of the project is proven to be successful and therefore it can be further enhanced to classify different types of images such as section details and even texts.

9. Conclusion

In conclusion, although AI may be portrayed as a threat to humanity (Bostrom, 2014), ML has a great impact in our everyday lives has potential in benefitting the design workplace. Through this research project, proving the possibility to automate image classification of plans and sections is the first step to the future, where further features such as text – digital and hand-written, section details and more complex designs are to be categorised in making a better search engine for the engineers/designers. For these features to be recognised in the classification, it requires further hyperparamater tweaking to suit the desired outcome. For the projec to be taken further, a streamlined model is aimed to be implemented into company/firm's systems, where files are automatically processed once received from clients. Moreover, this does not only apply to the built environment discipline, but can also be utilised by others that are in need for complex image classification.

Acknowledgements

I would like to acknowledge the guidance and support from Built Environment UNSW and Arup, family and friends.

References

Zendesk. (2018). A simple way to understand machine learning vs deep learning - Zendesk. [online] Available at: https://www.zendesk.com/blog/machine-learning-and-deep-learning/.

Basheer, I. and Hajmeer, M. (2000). Artificial neural networks: fundamentals, computing, design, and application. Journal of Microbiological Methods, 43(1), pp.3-31.

Sharma, A. (2017). Convolutional Neural Networks in Python. [online] DataCamp Community. Available at: https://www.datacamp.com/community/tutorials/convolutional-neural-networks-python [Accessed 2 Sep. 2018].

Gatys, L., Ecker, A. and Bethge, M. (2015). A Neural Algorithm of Artistic Style.

Li, H., Ellis, J., Zhang, L. and Chang, S. (2018). PatternNet: Visual Pattern Mining with Deep Neural Network.

Shah, D. (2018). AI, Machine Learning, & Deep Learning Explained in 5 Minutes. [online] Becoming Human: Artificial Intelligence Magazine. Available at: https://becominghuman.ai/ai-machine-learning-deep-learning-explained-in-5-minutes-b88b6ee65846 [Accessed 1 Sep. 2018].

Bostrom, N. (2017). Superintelligence. Oxford: Oxford University Press.

Ayers, R. (2018). The future of Artificial Intelligence: 6 ways it will impact everyday life. [online] Big Data Made Simple. Available at: https://bigdata-madesimple.com/the-future-of-artificial-intelligence-6-ways-it-will-impact-everyday-life/ [Accessed Oct. 2018].

Phelan (2016). Designing with Machine Learning. [online] WeWork. Available at: https://www.wework.com/blog/posts/designing-with-machine-learning [Accessed Nov. 2018].

Zulkifli, H. (2018). Understanding Learning Rates and How It Improves Performance in Deep Learning. [online] Towards Data Science. Available at: https://towardsdatascience.com/understanding-learning-rates-and-how-it-improves-performance-in-deep-learning-d0d4059c1c10