Abstract. With the 'digital turn', machines now have the intrinsic capacity to learn from big data in order to understand the intricacies of architectural form. This paper explores the research question: how can architectural form become machine computable? The research objective is to develop “Deep Architectural Archiving” (DAA), a new method devised to address this question. DAA consists of the combination of four distinct steps: (1) Data mining, (2) 3D Point cloud extraction, (3) Deep form learning, as well as (4) Form mapping and clustering. The paper discusses the DAA method using an extensive dataset of architecture competitions in Switzerland (with over 360+ architectural projects) as a case study resource. Machines learn the particularities of forms using 'architectural' point clouds as an opportune machine-learnable format. The result of this procedure is a multidimensional, spatialized, and machine-enabled clustering of forms that allows for the visualization of comparative relationships among form-correlated datasets that exceeds what the human eye can generally perceive. Such work is necessary to create a dedicated digital archive for enhancing the formal knowledge of architecture and enabling a better understanding of innovation, both of which provide architects a basis for developing effective architectural form in a post-carbon world.

Keywords. Artificial Intelligence; Deep Learning; Architectural Form; Architectural Competitions; Architectural Archive; 3D Dataset; SDG 11.

1. Introduction

The roles attributed to form and the attendant positions within the discipline of architecture have undergone notable transitions that, in effect, reframe the very status of form itself (Eisenman, 1999). The focus on the genesis of form shifted over time as it was overshadowed by various social, cultural, ecological, or political influences on architecture. However, with recent developments of deep learning via neural networks, machines now have the capacity to learn from unlimited volumes of data, thereby lending novel significance to the issue of form-making in design disciplines, while also
enabling architects to continue their disciplinary focus on form, albeit in the still-emergent interface between human and machine (Huang et al. 2021). This research explores the following overarching research question: how can architectural form become machine computable?

The transition to the machine age allows form to operate transversally in the environment where it embodies “all forms of ecology together, whether environmental, mental, or social” (Moussavi & Lopez, 2009). However, even with such innovative thinking, there are still limitations in developing applications of 3D machine learning (ML) models because of the complexity that arises from higher dimension data and the flexibility that decreases with different representations of 3D data such as point clouds, wireframes, voxels, and meshes (Zhang & Huang, 2021). Furthermore, the inconsistencies in methods, details, and resolution of data produced by different individuals also hinder the gathering of consistent 3D data required to compile a cohesive dataset (Stoter et al., 2020). Finally, compared to readily available 3D objects with distinct shapes, such as chairs or sofas, that have been widely used to develop the 3D ML techniques buildings have considerably more formal variables that must be quantified and categorized for archiving purposes. Moreover, such variables demand even greater efforts in compiling the respective architectural datasets.

To address such limitations, this paper introduces a new technique for 3D deep learning of architectural forms: The Deep Architectural Archiving or DAA. DAA employs four steps: (1) Data mining, (2) 3D Point cloud extraction, (3) Deep form learning, and (4) Form mapping and clustering. The paper discusses the procedure in detail and applies the technique to a competition dataset as a case study resource. The application of DAA to a competition dataset can contribute to the structuring of culture and building knowledge in a materialistic and intellectual way for the edification cultural heritage (Chupin et al., 2015). In addition, the dataset for this research is an archive of architectural competitions in Switzerland from 2009 to 2021 that consists of unexploited knowledge of formal innovation, much of which is often undervalued, if not ignored outright. This repertoire of architectural propositions provides the necessary and extensive data for closely examining understudied relationships, distinguishing morphological patterns between architectural forms across different competitions using deep learning, and creating an archive of architectural competition comprising a 3D dataset using ML as a formal indexing strategy.

2. Background

Due to the inherent complexity of 3D information and the computational power needed to learn 3D datasets, 3D machine learning techniques require the development of creative methods capable of achieving an appropriate format for the machine readability of 3D architecture. Several methods have been developed for machines to learn from different data formats such as multiple 2D images, voxels, point clouds, or wireframes.

Using a dataset of 2D images, Steinfeld et al. proposed a framework for how 2D ML can be synthesized with 3D generative architectural design (Steinfeld et al. 2019). The volumes of gabled bay houses are translated into multi-view, tiled projection.
images to feed into Generative Adversarial Networks (GAN), which generate a composition of elevations that can be constructed into a new form. Similarly, Zhang and Huang also created new forms from different building styles by using 2D images, a sequence of section cuts of 3D buildings, and the multiple layers of training networks such as StyleGAN, Waifu2X, and Pix2Pix (Zhang and Huang, 2021). Newton experimented with GANs using 3D voxels of massing models in downtown NYC by showcasing the results of 3-D I WGAN for encoding and generating new building forms (Newton 2019). In addition, Rodríguez et al. proposed using a 3D grid of wireframes of buildings and variational autoencoder (VAE) (Rodríguez et al. 2020).

Inasmuch as the previous approaches have demonstrated the efficacy of creative methods utilizing different formats of data with 3D ML algorithms for the field of architecture, the current research will deploy an unsupervised machine learning model called 3D Adversarial Autoencoder (3dAAE) using point clouds (Zamorski et al. 2020). The autoencoders are used for a discriminative task, where inputs are data and outputs are both latent points and reconstructed data, which allows for classification and representation learning for indexing and clustering. The model accepts direct input and generation of 3D point clouds composed of 2048 points for learning architectural form. Finally, the current research will be using the discriminative tasks of 3dAAE by using 3D architectural data and different point cloud strategies to develop a method for creating a machine-learned archive of architectural forms.

3. The DAA Method

The DAA method explores a framework of unsupervised machine learning of architectural forms using optimized point cloud representations to better understand the formal relationships among the 3D datasets. Figure 1 demonstrates a diagram of the proposed methodology, which includes (1) collecting data to create a 3D dataset of architectural forms, (2) translating the representation of data into machine learnable formats, (3) experimenting with ML to get the latent representation for organizing data, and (4) visualizing the training results using ML that allows humans to observe organizational capacity.

![Figure 1. DAA, the proposed method for machine learning of architectural forms.](image-url)
3.1. DATA MINING

The research utilizes the results of the well-organized Swiss architecture competition system as the source of the data. Specifically, the distinctively abstract white plaster massing models, required within the detailed procedural guidelines by the Swiss Society of Engineers and Architects (SIA), provide a key reference point and the primary source of data for the current research on form as they display essential and unadorned qualities of an architectural form in a given context, while also underscoring the importance attributed to form as an autonomous parameter during the evaluation process. Jury reports containing site plans, floor plans, sections, elevation drawings, and model photos from the web-based competition databases of Konkurado and Espazium are collected to reconstruct the white massing models in digital format. The 3D volumes are manually modelled to maintain accuracy and achieve structural consistency of the 3D models. Using Rhino 3D, each 3D volume was generated via an algorithmic approach. The first step was to scale the plans, sections, and elevations to 1:1 and position them in 3D space. The second step was to operate a series of extrusion and Boolean operations for constructing the volumes. Finally, in the third step, photographs of white massing models were used to verify any missing information from the drawings or complete the 3D modelling. Initially modelled using NURBS geometry, the models and can be subsequently transformed into different data formats that are required by various ML algorithms.

The 3D models are developed with respect to their sizes in real-world dimensions and orientation in context. The results from 63 competitions of schools that ranged from 3 to 10 in project ranking are translated into 366 3D volumes from the past fifteen years. The 3D volumes are then used to generate a cohesive dataset for exploring architectural forms. Figure 2 illustrates the dataset in order of ranking on the vertical axis and chronological order on the horizontal axis.

Figure 2. 3D dataset of 366 models from 63 school competitions in Switzerland arranged in chronological order on the x-axis and rankings on the y-axis.
3.2. 3D POINT CLOUD EXTRACTION

This step entails the preparation of a dataset, with 3D volumes transformed into an optimal representation of the architectural form using 'point clouds' as a machine learnable data format. Rather than a typical uniformly distributed point cloud, an 'architecturalized' point cloud is considered by optimally distributing 2048 points to represent the 3D architectural information for ML training. 'Architectural' point clouds attempt to structure the points in a meaningful architectural representation that both machines and humans can understand. As seen in Figure 3, inspired by the idea of the ‘G-codes’ (or geometric codes), a tool path produced by slicing a model for 3D printing is used to distribute the points. The points are evenly spaced out along the lines of the slicing paths to create consistency in the structure of point clouds. Using the G-code to extract the points allows different data formats of 3D volumes to be translated into a consistent dataset. The technique also normalizes the location of the points along the z-axis for each point cloud and reduces a degree of freedom for machine learning.

Since there is no architectural scale in ML, the question of how 3D models are generalized from different sizes of buildings must be addressed. To develop the dataset, each model was scaled relative to the largest model to fit into a bounding box of 1×1×1 unit. In addition, a custom Grasshopper definition was developed using Rhino 3D to automate the process of transforming the input geometry into 'architectural' point clouds or uniformly distributed points. For the ML experiment, two-point cloud generation methods were used for developing two datasets: standard uniformly distributed point clouds and 'architectural point clouds.'
3.3. DEEP FORM LEARNING

Using the two datasets, the 3dAAE model, an unsupervised machine learning algorithm (Zamorski et al. 2020), is trained with 15,000 epochs, a learning rate of 0.0001, a batch size of 5, and the latent space dimension of 128. The training results of encoding and decoding 3D data are evaluated qualitatively by comparing the reconstruction of the point cloud to its sampled point cloud from the original architectural form. The success of the results is demonstrated in the regeneration of the point clouds that are remarkably similar to the original architectural form. Figure 4 shows the progression of the reconstruction of the point clouds at different iterations. In this example, the network shows successful learning of data composed of two rectangular volumes. As seen at 1,000 epochs, the point clouds split in two. After 10,000 epochs, the reconstructed data show stability in their point-cloud structure.

![Figure 4. Sample of training progress at different epochs.](image)

Fig. 5 shows notable differences between using uniformly distributed point clouds and architectural point clouds in ML training. Reconstructed data using uniform point clouds show randomly distributed points with a single region of a dense cluster of points within each volume. Although overall shapes are captured from the input data, there is still a lack of architectural details. As for the architectural point clouds, reconstructed data captures the general shapes and formal characteristics of architecture, such as the details of the surfaces along the volume. Compared to the single region of concentrated points when using uniformly distributed point clouds, reconstructed architectural point clouds resemble a magnetic field with the interplay of two concentrated regions of points. Figure 5 illustrates the results of comparing reconstructed and input data of uniformly distributed point clouds on the left and architectural point clouds on right at 10,000 epochs. As seen in the input data of architectural point clouds, the horizontal planes do not contain the points due to their z-direction on the slicing of volumes. This leads to a higher concentration of points around certain elements in the buildings for a clear representation of the reconstructed point clouds. For both point cloud structures, the smaller models captured greater levels of details than the larger ones due to their compactness of the density of points caused by the limited number of points for representing different sizes of 3D models.
3.4. FORM MAPPING AND CLUSTERING

During the machine learning training, the data is compressed to 128-dimensional encoded representations and visualized into 2D space using t-Distributed Stochastic Neighbor Embedding (t-SNE) (Maaten and Hinton 2008). The translation allows for the interpretation of the training results, insofar as the original 3D models are projected according to their encodings in the latent space. The spatialized arrangement of the machine-read relationships offers a more fundamental understanding of correlations across different formal groups expressed by the affinities among the forms. In this process, the t-SNE technique is used to map out the volumetric models at the encoded data points at every 1,000 epochs to evaluate the performance of ML using a qualitative evaluation of human interpretation by observing the placements of architectural forms. Rather than the reconstructed point clouds, the original volumetric models are used as the visual representations of the encoded data points. For example, Figure 6 shows the results of t-SNE plots at 9,000 epochs of previously mentioned experiments using two different point cloud datasets for ML training.
As for comparing the results between two different point cloud structures, the example shows subtle differences in the overall mapping of the 3D forms. The performance of an ‘architectural’ point cloud is compatible with the performance of standard point clouds. Furthermore, the positioning of different clusters of formal groups in the architecturalized point clouds even shows smooth transition in the arrangement of the volumes. For example, the clusters of linear bar buildings in different directions appear to be closer to each other in the mapping of volumes. For further experimentation, one can also imagine orienting the 3D models in the same direction to exclusively examine shapes as the main comparative feature. However, t-SNE showed a pitfall after 9,000 epochs. It showed a collapsing behaviour with merged clusters, whereby it was no longer possible to find a correlation between the encoded models (Wattenberg et al. 2016).

**Figure 6.** t-SNE plots of 3D models showing training results from two datasets at 9,000 epochs.

### 4. Results and Discussion

DAA is proposed as a viable approach for making architectural form machine-computable and for enabling humans to analyse and interpret the machine understanding and learning of 3D architectural forms. Since it is not quantitatively evident how machines understand forms, the visualization of latent space for the qualitative measure is essential for examining the performance of encoding and decoding processes of networks at different stages of training. As shown in Figure 7, it is clear that the network understands the formal properties of the 3D models at multiple scales. In the overall view of the latent space, the models are positioned in a gradient from large to small buildings. The orientation of the buildings is reflected in the mapping, which creates a smooth change in direction between the models like a vector field. In addition, different shapes are clustered together. For example, linear buildings are placed in the perimeter of the latent space, and cubic-shaped buildings are placed towards the central area. Zoom-ins on the right side of Figure 7 reveal details such as voids, courtyards, the complexity of roof surfaces. Finally, facades are reflected in the proximities between the 3D models.
While the proposed method is promising, there are limitations with the DAA method worth noting: (1) time-intensive manual labour in translating 2D architectural drawings into 3D data. (2) a lack of data that makes it difficult to identify the subtle differences of the mapping results of learned forms. To address these limitations, further experimentation will be undertaken within each step of the DAA method. For example, an increase in the number of 3D datasets from the Swiss competition archives is currently in progress. Optimization of the number of points and structure of the point clouds will also be investigated. In addition, different hyperparameters, such as the number of dimensions, learning rates, and batch sizes, can be further explored to optimize the training performance. Finally, clustering algorithms, such as K-means could be used to identify machine categorizations of encoded models.

5. Conclusion and Future Work

This research develops a 3D dataset drawn from the architecture competitions in Switzerland to examine how machines can understand the particularities of architectural forms and evaluate comparative relationships among form-correlated datasets that far exceed human comprehension in their scale and complexity. Using DAA, the findings show promising results in reconstructing architectural forms by optimizing data structure for the training of networks using a relatively small dataset composed of uniform and 'architectural' point clouds. The visualizations of the results with t-SNE mapping demonstrate how formal properties such as geometrical features, scale, and orientation can be identified and how hidden connections can be discovered.

Figure 7. t-SNE plot of machine-learned forms using ‘architectural’ point clouds at 9,000 epochs with close-ups of clustering of different formal groups.
The trans-scalar view of the forms from different competitions enables a more precise examination of understudied relationships among architectural forms.

In view of future applications, DDA aspires to create a novel multi-dimensional and dynamic digital archive that facilitates a machine-enabled reading of architectural forms. It aims to advance our understanding of the inherent morphological patterns within the collection of 3D data and offer insights to architects in their formal exploration throughout the design process. Such an archive could index not only the formal properties but also preserve building data such as program, area calculations, cost estimates, energy efficiency, time factor, and jury rankings from design phases contributing towards sustainable building culture. It is also foreseeable that such knowledge could enable a detailed understanding of the distinct contextual, ecological, and cultural relevance of specific architectural forms in the post-carbon world.

References


