ABSTRACT
Architecture does not exist in a vacuum. Its cultural, conceptual, and aesthetic agendas are constantly influenced by other visual and artistic disciplines ranging from film, photography, painting and sculpture to fashion, graphic and industrial design. The formal qualities of the cultural zeitgeist are perpetually influencing contemporary architectural aesthetics. In this paper, we aim to introduce a radical yet methodical approach toward regulating the relationship between human agency and computational form-making by using Machine Learning (ML) as a conceptual design tool for interdisciplinary collaboration and engagement. Through the use of a highly calibrated and customized ML systems that can classify and iterate stylistic approaches that exist outside the disciplinary boundaries of architecture, the technique allows for machine intelligence to design, coordinate, randomize, and iterate external formal and aesthetic qualities as they relate to pattern, color, proportion, hierarchy, and formal language. The human engagement in this design process is limited to the initial curation of input data in the form of image repositories of non-architectural disciplines that the Machine Learning system can extrapolate from, and consequently in regulating and choosing from the iterations of images the Artificial Neural Networks are capable of producing. In this process the architect becomes a curator that samples and streamlines external cultural influences while regulating their significance and weight in the final design. By questioning the notion of human agency in the design process and providing creative license to Artificial Intelligence in the conceptual design phase, we aim to develop a novel approach toward human-machine collaboration that rejects traditional notions of disciplinary autonomy and streamlines the influence of external aesthetic disciplines on contemporary architectural production.
**INTRODUCTION: DISCIPLINARY AUTONOMY AND CULTURAL ENGAGEMENT IN THE POSTDIGITAL ERA**

Since the commonplace integration of digital design and fabrication tools into architectural processes, much of the disciplinary effort focused on using computation as a tool for automating the production and fabrication of complex and variable formal systems efficiently and economically. The conceptual objective behind these schemes was to liberate architectural production from the Modernist norms of repetition and mass production, allowing for more diversity of formal languages to emerge. Nevertheless, in most of these parametrically-driven design and fabrication processes, the role of computation has traditionally been linear and subservient to the architect’s creativity and disciplinary zeitgeist. The objective of the computational process was to iterate on and automate the production of the architect’s sketch. Architecture was viewed as an autonomous discipline with its own set of tools and priorities, engaging with the culture of technology almost exclusively from its own disciplinary lens. Its affiliation with outside disciplines has largely been limited to social sciences and philosophy in order to mine for metaphorical formal models to justify many of its seemingly novel aesthetic agendas (Jarzombek 2016). This tool-centric approach allowed for architectural discourse to create many unique formal agendas, yet its dependency to the capabilities of digital modeling tools eventually limited its output and created self-similar results. During this time, the computational process operated as a black box that was not open to outside cultural influences, incapable of making vital intellectual, visual, aesthetic and conceptual connections to artistic disciplines other than itself. This isolation further limited its future potential and jeopardized its cultural relevance.

**CONTEMPORARY GENERATIVE DESIGN TOOLS: SIMULATION BASED PROCEDURAL MODELING VS. MACHINE LEARNING AND ARTIFICIAL NEURAL NETWORKS**

Since the inception of complex simulation tools that can generate and project formal schemes in a non-linear and meta-parametric fashion, we are able to duplicate the behavior of natural phenomena and harness the intelligence of natural systems as design models. Contemporary procedural modeling and animation software can accurately simulate the behavior of these natural systems, such as physics-based particle systems, fluid dynamics, agent-based systems and so on, and link them with the formal agenda of the architect’s choosing (Lienhard et al. 2017). Nevertheless in this type of predictive intelligence, the software provides still aims to create a subservient hierarchy between the tool and its user. Its capabilities as a tool for form-finding are predictable and linear. Contemporary ML tools however, based on artificial neural networks, allow for a much higher level of generative potential, wielding a level of autonomy and unpredictability never before seen in computational design tools. This type of new design intelligence—one that is capable of producing its own interpretations as it interpolates between predetermined computational objectives—provides an immense amount of creative potential (Karras, Laine and Aila 2018).

The difference between simulation-based generative modeling and ML-based projective interpolation highlights the differentiation between the performative potential of machine intelligence and its ability to automate form-making processes. Contemporary use of ML, particularly in the context of machine vision and sensor systems, are primarily focused on their interactive potential and their ability to respond to outside contexts such as the environment, human occupation, comfort, and psychology. Through cyberphysical architectural systems—that combine static components, responsive robotic components and media applications controlled by machine vision and sensor networks—we can create programmed behaviors for architecture to interactively and autonomously respond to environmental and occupational contexts (Ozel 2016). This approach yields productive results towards the understanding of human presence as it relates to architectural form and motion, but has not related to alternative spatial languages that can be deployed and explored further in order to pose alternatives to familiar models of computational form-making, which are heavily constricted by the limitations of parametric design tools. In many cases, questions of enclosure—as it relates to motion, novel material science experiments, and variable formal modulations to strike a balance between static and dynamic qualities of space—are explored through investigating the historic evolution of such forms in architectural and industrial design and are independent from their potential to interact with the world of culture and aesthetics around them.

In order to overcome this problem, this paper will focus on the relationship between human agency in design discourse and computational iteration by using Machine Learning as a generative concept design tool. We will be introducing highly customized Machine Learning tools in order to classify and iterate stylistic approaches that exist outside the discipline of architecture to allow for ML to design, coordinate, randomize, and iterate qualities as they relate to pattern, color, proportion, hierarchy, and formal language. The human engagement in this design process will be limited to the initial curation of input data.
that the ML system can learn from, and also in regulating and choosing the iterations as the final outputs of two-dimensional images such systems are capable of producing. Considering that contemporary ML systems are limited in their generative abilities to only provide 2D images, an additional computational mediation process in the form of agent-based systems and procedural modeling is deployed thereafter in order convert these images into 3D geometry, further automating and streamlining formal production. By questioning the notion of human agency in the design process and providing creative license to Artificial Intelligence in the conceptual design phase, we aim to develop a novel approach toward human-machine collaboration that rejects traditional notions of disciplinary autonomy and streamlines the influence of external aesthetic disciplines on contemporary architectural production.

PROCESS
Documenting Selected Architectural Landmarks
In order to limit the architectural scope of our experimentation, we decided to pick five landmark buildings in downtown Los Angeles as our sites of experimentation. Through documenting these selected civic buildings and their urban presence in downtown Los Angeles, we aimed to bypass discussions regarding urban mass, adjacency and other architectural and urban concerns that fall beyond the scope of the proposed design method. The existing massing of the selected buildings and their inherent connection to their context served as vessels for our machine learning operations. In order to achieve an accurate and comprehensive documentation of these buildings, we used existing drone footages from various angles as the basis for our generative Machine Learning process (Figure 2).

---

2. Drone footage of architectural landmarks in downtown Los Angeles
Precedent Research and Selection of Interdisciplinary Aesthetic Influences

Our main cultural objective is to look outside the discipline of architecture to mine for inspiration for new aesthetic paradigms. Looking into the world of film, photography, painting, sculpture, fashion, and industrial design, we asked our students at the Technology Studio at UCLA’s Department of Architecture and Urban Design, IDEAS Program, to create a detailed repository of images from a particular body of work by an artist or designer who is alive and influential. This library of images later on served as the basis for “Style-Transfer” from the image into the massing of the building as documented through the drone footage (Figure 3).

METHOD

Artistic Style-Transfer with Convolutional Neural Networks on Architectural Drone Footage

We developed a technical pipeline to use artistic Style-Transfer operations where the influence image is classified as input, the drone footage divided into single frames as target image, and the new generated image as output content (Figure 4). Based on that workflow, we asked our students at the Technology Studio to apply the technical pipeline of artistic Style-Transfer on Drone Footages of the five landmark buildings in downtown Los Angeles. The Style-Transfer algorithm that we utilized in this workflow is a machine learning algorithm using Convolutional Neural Networks (CNN) that is executed through Google’s TensorFlow and Nvidia’s cuDNN API (Smith 2016). The Style-Transfer algorithm combines a target image with one or multiple style input images in order to generate a new output image. The new artificially generated images respect semantic features of the target and the style input image simultaneously. In this process of Style Transfer, the target images, the content and the style of the input images are separated and combined in multiple layers of the neural network in order to extract and deploy specific features in them. Each of these layers are small computational units in the CNN that processes visual information and classifies them (Gatys, Ecker and Bethge, 2015). Through this classification, we are able to distinguish patterns in style and content images and map features from one to the other. This process is considered as part of a texture transfer problem in computer vision (CV) by using machine learning and neural networks. We are particularly interested in using this technique on content images (output) as a resultant of architectural elements (target, drone footage) and stylize them with artistic images (input) since this process allows for maintaining the architectural characteristics of certain images. For example, the system is capable of recognizing elements such as window, door, roof and other architectural features and is able to maintain their general

Sougwen Chung  Gerhard Richter  Tony Cragg  Alexander McQueen

Yayoi Kusama  Yayoi Kusama  Universal Everything  Meta Heaven

Takashi Murakami  Olivier Rousteing  The Haas Brothers  Bridget Riley

3 Interdisciplinary aesthetic influences from visual art, sculpture, painting, fashion and industrial design
morphology while transferring the semantic qualities of the input style influence image. This technical setup creates a highly accurate architectural definition that consequently respects the initial functional parameters of the building.

Relevance of the Style Image
In our workflow and research, the quality of the input style image is particularly important to control and steer the final result. We discovered that style images with clear patterns and defined color palette have a much higher success rate in being transferred into content images with an architectural subject. The patterns in the style image can change in scale and density as long as they follow the same set of distinguishable rules. If the style influence is blurred with no clear regions or boundaries, the result is equally undefined and only the color palette is being transferred. This is based on the object detection capabilities of the CNN and the pretrained VGG-Network (Gatys, Bethge et al. 2016). In this regard, we introduced an additional step to the process by editing the style and input images. Through experimenting with contrast ratios, saturations, and color values of the target image and input image, we could calibrate the success rate of the output image. Through this process, we can increase the strength of pattern recognition at the level of the CNN and curate the final result towards a particular direction. Further, we discovered that the pixel ratio of the target and as well as the input style image can influence the final result. The Style Transfer algorithm natively transforms the style image into the pixel ratio of the target image. If the format and pixel count of the input style image is significantly different than the target image, the style image will be scaled and interpolated. This operation within the algorithm can yield to undesirable results in terms of resolution and sharpness. Therefore, the inbetween image manipulation steps we introduced to the style image prior to executing the transfer heavily influenced the success rate of the generated result.

Style Weight and Multiple Input Styles
Despite the style (input) image we can control the number of style influences and their weights within the Style-Transfer algorithm. This gives us the possibility to adjust the trade-off between the target image and style influence. By simply applying more than one style image, the CNN will try to keep the semantic features in all images but generates a hybrid or a pastiche between the target and style input images (Smith 2016). Further experimentation with different style weights allowed us to prioritize certain influences in the process of texture transfer that is aware of all the semantic features and patterns in an image composition. In this method, each input image that is used as a style reference is given a particular weight. This method alters
the final outcome dramatically but is always applied globally to the whole content image. In order to apply multiple style images to selected regions on the content image, we use paired image masks that assign a particular style influence to a specific boundary on the content image (Figure 5). This method is also referred to as semantic image segmentation in Computer Vision (CV) and adds an additional layer of control to the process (Smith 2016).

Feature Extraction of Architectural Elements
In our process, we were particularly concerned about the detection and recognition of architectural elements in 2D imagery and how formal characteristics are transferred through artistic Style-Transfer to these elements. In our artistic Style-Transfer process, we used the VGG-19 network for object recognition and feature extraction. This deep convolutional neural network was developed and trained by Oxford’s Visual Geometry Group (VGG) (Simonyan and Zisserman 2014). This pre-trained neural network has 1619 layers and was trained on vast datasets that are not specific to architecture and the built environment. Nevertheless, the neural net by VGG has proven to be very effective on a general basis of object and feature detection for various datasets (Simonyan and Zisserman 2014). Also in our process, the VGG-19 network has proven to be very effective in detecting and labeling architectural elements as features from the target image. In order to create a feature map from the style images and the target image, we reconstructed the input image from layers ‘conv1 1’, ‘conv2 1’, ‘conv3 1’, ‘conv4 1’ and ‘conv5 1’ of the original VGG-Network (Gatys, Bethge et al. 2016). We discovered that through the use of the VGG-19, architectural elements such as windows, doors, facade patterns, patios, setbacks and architectural ornaments are very well preserved as features and patterns in the output image. This yields to precise results while transferring the feature map of a style input image to the target image. We realized that in regards to the architectural elements, contrast and light conditions in the target image are crucial to extracting geometric figures such as windows and doors as rectangles or facade patterns as features for the Style-Transfer. But despite the pure feature extraction of architectural elements through the use of the VGG-19 network, we also discovered that the CNN is capable of detecting and accounting for factors such as scale, perspective and orientation of architectural surfaces and patterns. Localized structures and patterns are very well perceived and rendered in the style of the reference images. This was a crucial factor for the success of Style-Transfer onto existing architectural buildings and elements (Figure 6).
In our pipeline we looked for methods of translating 2D images into 3D geometry with procedural 3D modeling techniques. Procedural modeling is a method to generate and describe 3D geometry through a set of rules. With this method, the geometry is constructed on runtime and can be described through a series of computational instructions (Ullrich, Schinko and Fellner 2010). We utilized this method in Houdini, Cinema 4D, Maya, and Rhino's Grasshopper with our students at UCLA in order to establish a direct link from the process of style transfer to 3D geometry translation (Figure 7). More specifically we automatically linked procedural modeling techniques and mesh operations such as extrusion, loft, height field, tessellation, sweep, revolve and deform to 2D imagery. The visual content of the style transferred images is extracted by the procedural modeling software and translated into architectural 3D models. During this process we again utilized low level CV operations within the procedural modeling workflow such as region detection, image segmentation and feature extraction. This allows us to automate the translation from 2D images to 3D geometry and create a more coherent link for human-machine collaboration.

Since we are applying the Style-Transfer to architectural imagery and translating the 2D images to 3D models, we are interested in capturing architecture and a building from more than one elevation. Therefore, we used drone footage as our content images in order to get all elevations from a building and its context. In order to achieve such comprehensive documentation, we applied the Style-Transfer on a video sequence that was rendered in full HD with a Nvidia 1080 Ti graphics card with 11GB memory. In order to maintain a stable and continuous Style-Transfer on multiple frames we used latest loss functions and initializations for Style-Transfer with CNN (Ruder, Dosovitskiy and Brox 2016). Once we got a stable video sequence that is stylized through the reference images, we were able to reconstruct and generate the 3D model of the building with various techniques of procedural modeling and photogrammetry. In our particular pipeline from 2D stylized images to 3D geometry, we used the capabilities of Cinema 4D to reconstruct scenes and 3D models from video footage through photogrammetry (Kersten et al. 2015). This allows us to directly link the process of 3D generation to the video...
output of artistic Style-Transfer (Figure 8). Further, we were also able to use Cinema 4D’s built-in rotoscoping tools to later visualize and render the stylized 3D geometry composited back into the drone footage (Figure 9). This process allowed us to show the new style of the building in situ so that we could regulate the influence of Machine Learning on the urban performance and presence of the building.
CONCLUSION
The procedural modeling process of extracting a 3D model from the ML generated 2D image sequence requires qualitative decisions from the human designer in regards to what modeling techniques should be deployed and automated to get the most accurate results. Like in any design development process, any concept design goes through multiple steps of iteration with large teams, and no current technological process or algorithmic scheme is capable of delivering a complete design from start to finish, and the goal of the ML integration is not to fully automate the design process. Our technique challenges the traditional workflow of computational design in the sense that it creates a more dynamic exchange of agency between humans and machines. Previously, the human designer would come up with a concept sketch and would devise computational design methods to automate the design and production of the design concept through computation. In this new workflow, the Machine Intelligence comes up with the concept design by considering design priorities determined by the human designer, but the human designer will have to devise a combination of manual as well as procedural methods to convert the ML generated sketch from a 2D concept image into a 3D model. Therefore, our workflow exemplifies a speculative instance where Machine Intelligence engages with the design process in an intuitive and creative way for the first time, yet still relies on human collaboration for furthering the design.
Rendering of Style-Transferred US-Bank

Style-transferred and rotoscoped Drone Footage - US Bank

Rendering of Style-Transferred US-Bank

Style-transferred and rotoscoped Drone Footage - The Broad

Style-transferred and rotoscoped Drone Footage - LADWP

Style-transferred and rotoscoped Drone Footage - Griffith Observatory

Style-transferred and rotoscoped Drone Footage - City Hall DT Los Angeles

3D Geometry composited back into the Drone Footage
REFERENCES


Ullrich, Torsten, Christoph Schinko, and Dieter Fellner. 2010. ‘Procedural Modeling in Theory and Practice.’ Fraunhofer IGD.

IMAGE CREDITS
Figure 1, 9, 10: © Rui Ding, Allen Wang, Ying Huang, Hoodeen Hakimian

Figure 2: Drone Footage by: © Misha Larkin, Mingomatic, Evan Thomas, DJI Planet, Michael Garcia, Around The World 4K and James - BeBop

Figure 3: © Sougwen Chun (Systole, Diastole), Gerhard Richter (Abstraktes Bild), Tony Cragg (Ferryman) Photography: Stefan Kaiser. Alexander McQueen (The Horn of Plenty) Photography: Nick Waplington, Yayoi Kusama (Pumpkin), Yayoi Kusama (The Moving Moment When I Went to the Universe), Universal Everything (The Science Museum - Presence), Metahaven (Bold Tendencies), Takashi Murakami (Flowers Blooming in the World and the Land of Nirvana), Olivier Rousteing (Balmain Embroidered Mesh Yoke Dress), The Haas Brothers (Accretion), Bridget Riley (Descending)

Figures 4: © Lu Geng, Sarah Kim, Qin Yang Lu, Rui Ding, Allen Wang, Ying Huang, Hoodeen Hakimian, Kshama Swamy, Haocheng Dai, Yiran Zhou, Universal Everything (Nike FIT), Sougwen Chun (Systole, Diastole & Mimicry Sequence 03), Bridget Riley (Descending & Fission & Nataraja), Marc Quinn (Chelsea Flower Show) Photography: Prudence Cuming

Figures 5: © Lu Geng, Sarah Kim, Qin Yang Lu, Rui Ding, Allen Wang, Ying Huang, Hoodeen Hakimian, Kshama Swamy, Haocheng Dai, Yiran Zhou, Chuck Close (Self Portrait), Universal Everything (Nike FIT & AOL Brand), Sougwen Chun (Systole, Diastole), Bridget Riley (Descending & Fission), Olivier Rousteing (Balmain Embroidered Mesh Yoke Dress)

Figures 6: © Lu Geng, Sarah Kim, Qin Yang Lu, Rui Ding, Allen Wang, Ying Huang, Hoodeen Hakimian, Kshama Swamy, Haocheng Dai, Yiran Zhou, Universal Everything (Nike FIT), Sougwen Chun (Systole, Diastole), Bridget Riley (Descending & Fission)
Güvenç Özel is an architect, artist, and technologist. He is a Suprastudio Lead at UCLA A.UD IDEAS Program, and the principal of Özel Office. His work is at the intersection of architecture, technology, and media. His projects and experimental installations were exhibited in museums and galleries in the USA and Europe, including Istanbul Museum of Modern Art and The Saatchi Gallery in London. His recent design and research on 3D printing was awarded one of the top prizes at NASA’s 3D Printed Habitats Competition. His latest installation, Cypher, was sponsored by Google’s Artists and Machine Intelligence Program and debuted at SXSW. At UCLA IDEAS, his masters design studio research focuses on virtual reality, machine learning/AI, robotics, and sensing interfaces with support from leading companies such as Autodesk, Microsoft, Oculus and others.

Benjamin Ennemoser is an architect and researcher based in Los Angeles, California. He has received several research fellowships, published his work all over the world and taught several classes and workshops at the University of Innsbruck, University of Applied Arts Vienna, University of Fine Arts Vienna, and Yun-Tech University in Taiwan. He is a licensed architect in Italy, and since 2016 he has been a lecturer at UCLA Architecture and Urban Design. As an architect and researcher, his work is situated within the field of advanced technology as design speculation and focuses on computational design, digital fabrication, applied robotics, soft-robotics, VR & AR, machine learning, and interactive architectural systems.