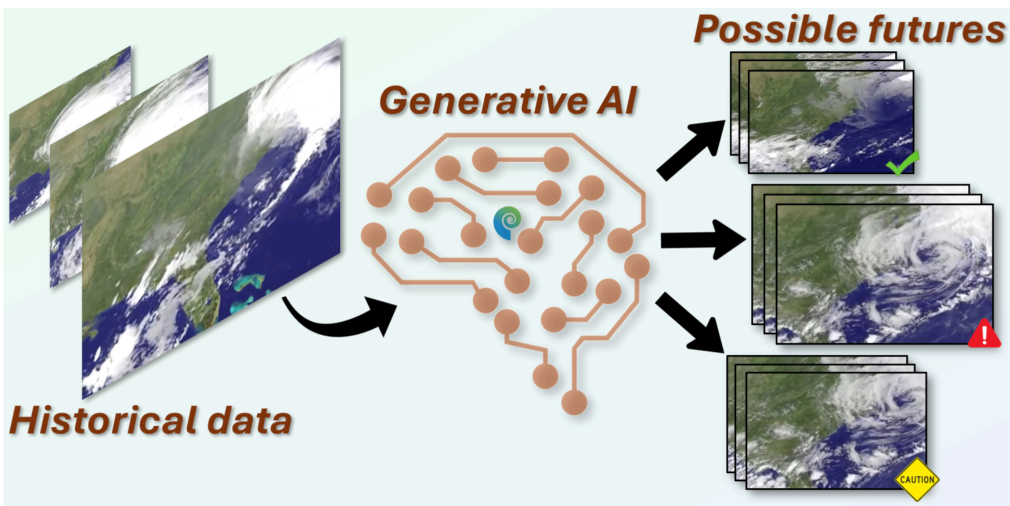
PHOTREK

Simulating Risky Worlds

for SingularityNET Deep Fund

R3-NEW-4 Milestone 4

Final Report



PHOTREK

Principal Investigator & Project Manager:

Kenric Nelson

Photrek, LLC

admin@photrek.io

11 October, 2024

Executive Summary

Photrek completed development for a Generative-AI (Gen-AI) capability enabling users to upload a sequence of imagery to a neural network (NN) model which learns the structural dynamics to forward-project a prospective outcome for use in risk management. Photrek hosted this capability on a cloud server, interfaced the NN to a Decentralized Application (DApp) for user interactivity, and submitted it to the Singularity Net AI Marketplace.

Photrek's key software developments in this milestone include (1) extending the VAE-RNN architecture to include deeper structures (up to 28-layer Residual Networks [Resnets]) for richer knowledge representation (to examine impact to preserving resolution of reconstructed imagery), (2) utilizing containerized software for microservices on cloud-hosted servers, developing and testing the DApp interface, and (3) evaluating model performance on representative time-sequenced imagery. Photrek analyzed VAE-RNN model's performance and determined that a "traditional VAE" – with a *single* latent layer performing generative sampling – may fundamentally inhibit high-resolution outputs from being generated. This conclusion was supported through a literature review of architectural variants (VAE-2, Nou-VAE, etc.). We observed deepening the architectures provided increased resolution of generated imagery and expect deeper structure with sampling across multiple latent layers will combine to produce higher-resolution reconstructions as our hosted model is updated.

Photrek's Simulating Risky Worlds (SRW) DApp was submitted to the SNET AI Marketplace. This DApp enables users to provide a 4-image sequence – 128x128 pixel color imagery – which is evaluated by the NN model and forward-projected into a new 4-image sequence. Subsequent calls may generate a new sequence so that users may create a family of prospective outcomes. Our DApp was designed to a baseline comprising representative dynamics (Celeb-A faces subjected to rotation) and an upgradeable NN model. We facilitate future updates within the DApp interface by providing users with a textual description of the present model structure and training data domain. These features ensure our DApp can be continually and responsively updated to align with customer needs.

Our DApp-hosted capabilities presently form an operational baseline – e.g., only rotational motion of Celeb-A faces are trained and – from the provided 4-image sequence – learned for use in forward-projection. However, the fact that imagery of this size is processed enables us to update the NN model hosted by our back-end service later. Our DApp interface includes features such as (1) indicating that back-end processing is underway when users invoke the service and (2) a textual summary of the type of model used by the back-end service. Users may reach out to Photrek's developer team should performance deficiencies be observed, thereby ensuring the DApp can be continually and responsively updated to align with customer needs.

Milestone Objectives

Develop a DApp interface for our model

Photrek developed a DApp enabling users to upload a sequence of 128x128 pixel 8-bit (256 levels per color channel) color imagery and receive an image sequence of prospective outcomes produced. We designed an interface to promote ease in users reviewing the provided inputs, staying apprised of the execution's progress, and understanding the outputs generated.

Users upload each of a 4-image sequence (in Portable Network Graphics [PNG] format) which are rendered within the DApp to ensure correctness. The DApp assumes that a constant inter-image time interval is associated with the dynamics responsible. Users are only able to invoke the service once all images have been provided (at which point the "Invoke" button is first enabled). Upon pressing "Invoke", the image inputs are provided to the back-end service which commences generation of a prospective outcome sequence. While the generative process is underway, an indicator message is provided so that users are made aware that the generative process may require up to 30 seconds. At conclusion, the 4 images generated are rendered within the DApp interface. Since we anticipate revisions as the backend service is trained with different architectures, we include a text field provided to the user from the backend model which describes the model architecture and training data domain used in the present invocation. This field ensures that users with questionable outcomes are provided information needed for subsequent contacting and expedient resolution of issues with Photrek developers.

Test and submit DApp to SNET AI Marketplace

Photrek tested the SRW DApp operation as hosted on the New England Research Cloud (NERC). Recall from previous milestone reports that NERC represents a cost-effective, collaborative, and readily-scalable option compared to hosting configurations used by Photrek in previous Singularity DApps. The testing phase also included tests ensuring that the NN model could perform inference with sufficient speed to meet the timeout constraints in which the DApp expected a response. To reduce server costs, Photrek implemented its trained NN model for operation on a CPU. This *inference* mode (on CPU) for trained models differs from the computationally-intensive *training* mode (on GPU) which produces the models used in inference. An additional benefit of Photrek's migration to NERC pertains to the prospect of using cost-effective access to modern server-grade

GPUs for training activities which we anticipate to facilitate expedient production of model upgrades in future developments. Cloud-based training ensures expedient operation with larger model architectures hosted on server-grade GPUs (typically 40-80 GB VRAM) than those available with on-premise consumer workstations (typically 16 GB VRAM).

Photrek developed the SRW DApp interface, tested it locally within the Singularity DApp “sandbox” environment, implemented incremental refinements, and finally submitted it to the Singularity Net AI Marketplace. Following submittal, we launched our ETCD cluster for payment processing and conducted tests with the AI Marketplace interface to ensure transactability.

Milestone Accomplishments

DApp interface for SRW was developed and tested (on a new platform – NERC)

We developed a DApp interface design to accommodate the user experience of customers seeking the forward-generation capabilities of SRW. We migrated DApp development and sandbox hosting activities to NERC. Component efforts – spanning generation of Protocol Buffer (protobuf) files, launching the NN-hosting backend service, launching the Singularity Net daemon (snetd), and launching the NodeJS-based sandbox – were each implemented in Docker containers to facilitate scalability and collaborative development.

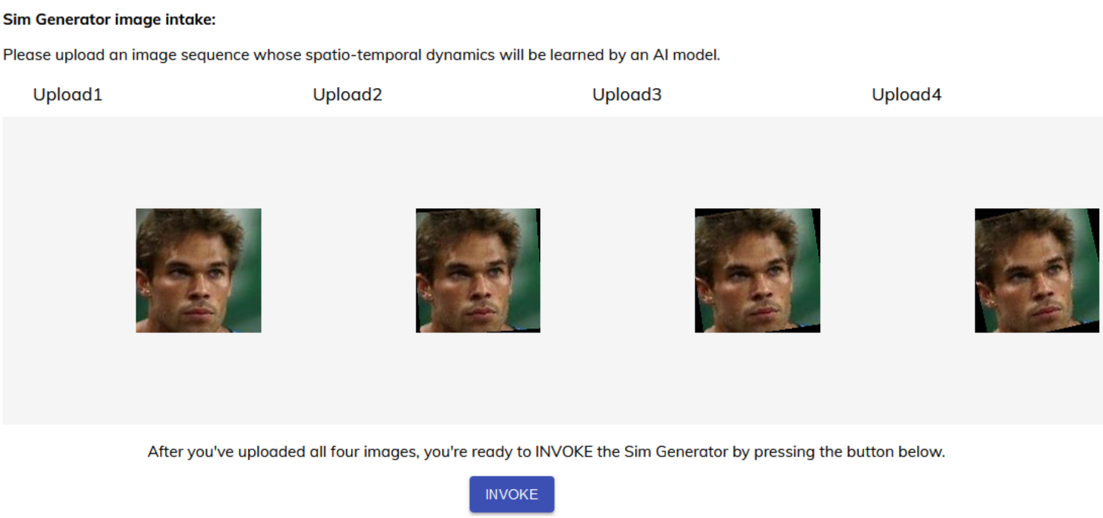


Figure 1: DApp interface enables users to review uploaded image sequence prior to service invocation. The "Invoke" button enables only upon successful upload of four PNG-formatted 128x128 color images.

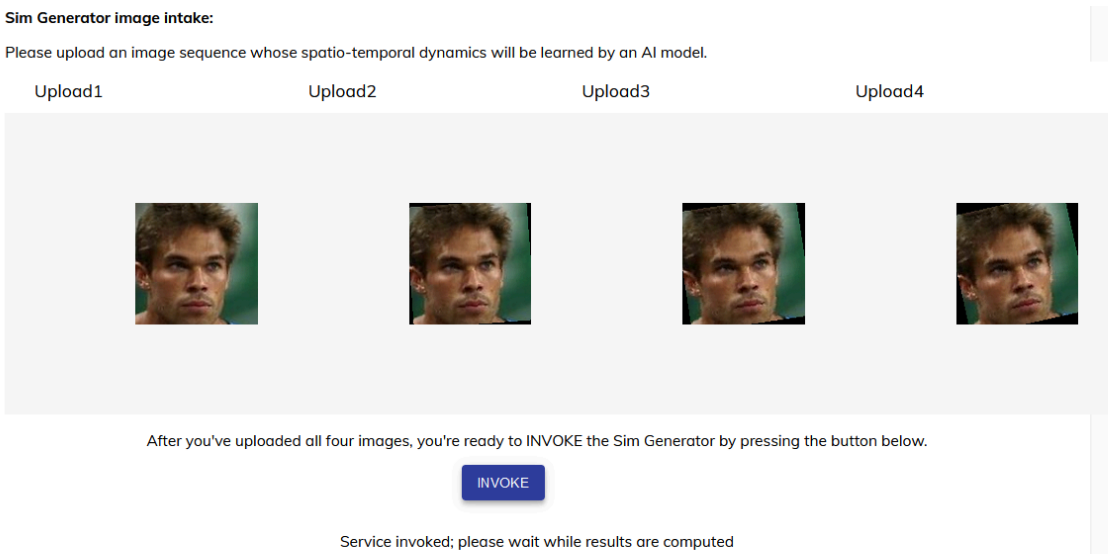


Figure 2: Upon invoking the service, users are notified that uploaded imagery has been provided to Photrek's back-end service and are presently being processed.

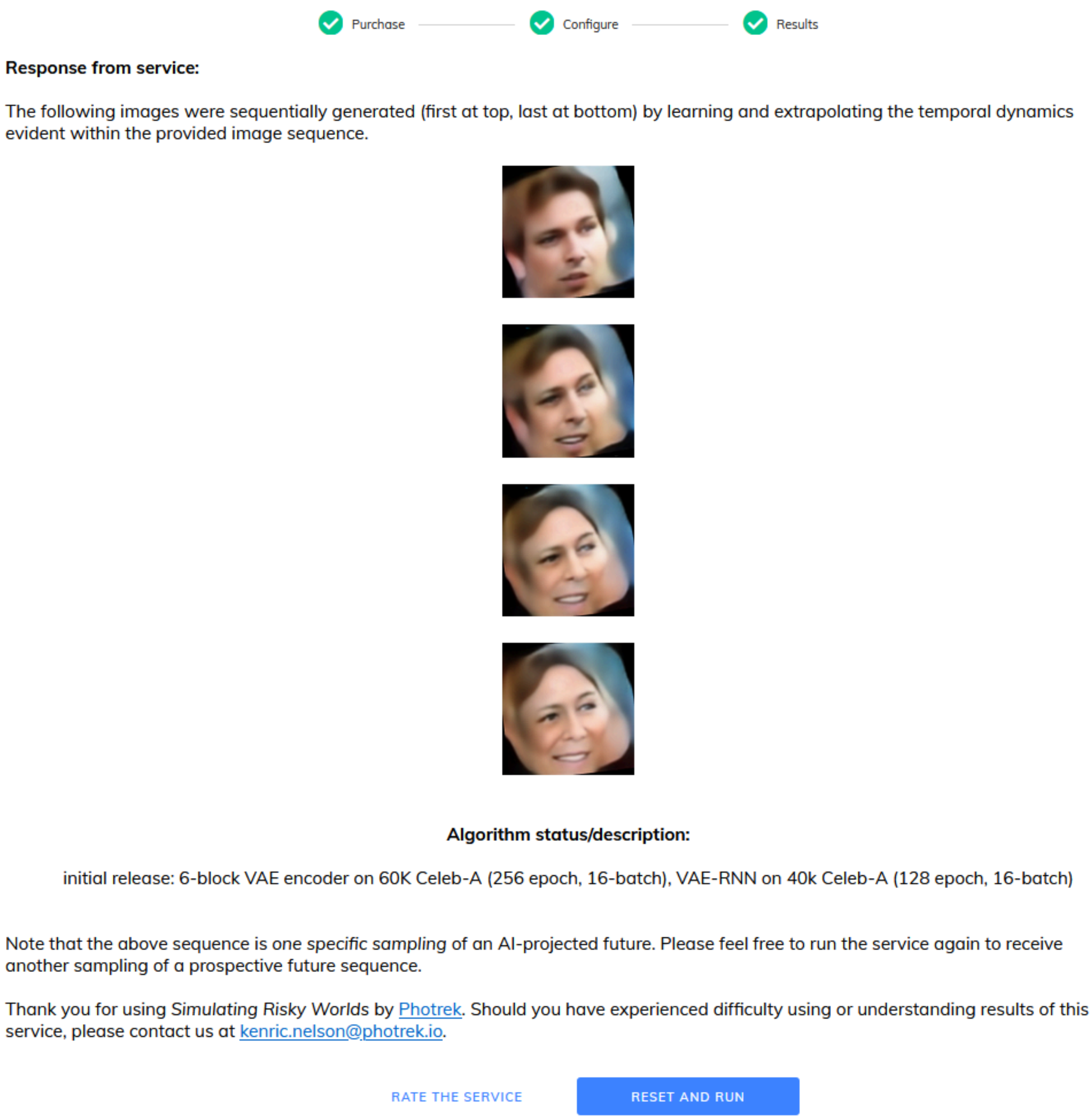


Figure 3: Post-invocation images are provided to the user along with a description of the back-end NN capabilities used in its generation and contact information for Photrek personnel to provide any needed assistance.

Model architecture refined to facilitate deeper structures and larger capacity

To explore the impact of *deeper* VAE model architectures – and hence increased knowledge representation *capacity* – on the quality of reconstructed imagery relative to prior efforts, Photrek parameterized the structuring of its VAE-RNN architecture (reported last milestone) and – in this milestone – used such to evaluate the impact of reconstructed image quality.

Photrek used a parametric form for the Resnet architecture in which the number of Resnet "blocks" (a term from the research literature) could be specified. Each block consists of two channels: (1) a channel for the residual connection, and (2) a channel comprising the following elements in sequence: a convolution layer, a batch-normalization layer, a second convolution layer, a second batch-normalization layer, and a final activation layer. Photrek configured the architecture construction methods so that 4-, 5-, and 6-block structures could be specified as a single configuration parameter so that training studies could iteratively examine resolution's impact to deeper structure.

Model performance and parameterization was evaluated

Photrek iteratively reconfigured the network (adding depth through additional blocks and increasing the number of neurons comprising each layer of the block) and evaluated the resultant quality of reconstructed imagery. We found that deepening the network structure afforded a noticeable mitigation of the "blurring" resolution degradation present in earlier work (both (1) 2 convolutional layers and (2) 3- and 4-block Resnets). The resulting resolution continued to exhibit observable resolution degradations for architecture depths up to 6 blocks. Further examination (including literature reviews) suggested this behavior is attributable to VAE structures in which feature sampling is performed at a *single* layer (typically the latent-most layer output from the VAE encoder) while VAE approaches using sampling across *multiple* latent layers may increase resolution of reconstructed output. Additional suggestions from the literature to increase resolution of VAE-generated imagery involve modified loss-function constructions. We anticipate exploring these approaches as we continue upgrading and training new models.

Also, in this milestone, we migrated from the 28x28 grayscale imagery of the MNIST dataset to the 128x128 center-cropped images of the Celeb-A dataset. The Celeb-A dataset enables a richer set of resolution criteria for evaluating VAE-RNN performance. Indeed, the spatial detail of the Celeb-A dataset was sufficient to reveal aspects of reconstruction quality so that our VAE model's structure could be iteratively reconfigured in studies seeking to preserve detail.

Even with the present model's constraints on reconstructable spatial detail, the VAE-RNN presents valuable forward-projecting potential for applications in which fine-detail may be of secondary importance. As an example, meteorological reflectivity maps from weather radars may exhibit broader spatial structure compared to human faces. These weather maps may be used in the SRW service's present capabilities to forecast prospective locations of storm fronts and concerning weather conditions.

NN training configuration established to use cloud-based computing assets

Training configurations to-date have exclusively utilized on-premise assets such as developers' individual workstations. Such assets have constraints on the available memory for the GPU (VRAM) which prohibit use of larger capacity models and the extent of data parallelism used in training (thereby increasing training duration). In this milestone, Photrek ensured GPU-enabled instances on the NERC cloud environment could be readily procured and used for training. At this time, the NERC GPU instances include access to nodes with 40 GB NVIDIA A100 GPUs. We ensured these GPU instances could be readily provisioned and utilized to expedite training requirements in subsequent NN training and evaluation.

SRW DApp was submitted to AI Marketplace

Photrek developed a DApp interface for the SRW service, tested DApp's operation in the Singularity Net "sandbox" environment, iteratively modified the NN-hosting back-end code (with associated revisions to the DApp user interface), and finally submitted the DApp to the Singularity Net AI Marketplace. Following submittal, we confirmed the model would operate as intended (including processing transactions) through tests in which Photrek's ETCD cluster was operational alongside the Singularity Net daemon and back-end service. Photrek also provided auxiliary components such as an overview graphic and textual descriptions to guide users on the intent and operational flow of the SRW service.

Budget & Schedule

Milestone	Description	Budget	Status
1	Contract Signing & Management Reserve	$ 2,150	Submitted & Paid
2	Dynamic CVAE	$ 18,043	Submitted Feb 07 & Paid
3	Upgrade resolution of SRW	$ 8,215	Submitted May 02
4	Onboarding of SRW	$ 8,752	Submitted Oct 30
5	Hosting Costs	$ 12,320	Planned over 6 months
Total		$ 49,480	

Future Plans & Change Notifications

Photrek anticipates hosting the DApp, aggregating user feedback, and continuing to train more capable back-end NN models as we seek to make the SRW service more useful and attractive to a variety of desired use-cases.

Photrek is also taking steps to socialize the SRW within the technical community. This includes an August 2024 in-person demonstration of its operation made at the AGI-24 conference. Photrek looks forward to future opportunities to showcase the advantages of its services – and the SingularityNet platform – in democratizing access to capable and useful AI services.

Based on user feedback, we may consider augmentations to the DApp interface in which the user provides information supplemental to the provided imagery in order to condition on the type of risk exposure sought in the generated imagery. These efforts will follow from the risk metric work in parallel grant and publication efforts (discussed in previous milestone reports).

Photrek graciously appreciates the Singularity Net team's assistance and support throughout the execution of this effort. We look forward to working on exciting and useful developments in future collaborations.

References

Cao, S., Li, J., Nelson, K. P., & Kon, M. A. (2022). Coupled VAE: Improved Accuracy and Robustness of a Variational Autoencoder. Entropy, 24(3). https://doi.org/10.3390/e24030423

Cheng, Ting-Yun, Marc Huertas-Company, Brant Robertson, Nesar Ramachandra, Francois Lanusse, and Claire Guilloteau. 2019. "Vector Quantized Variational AutoEncoder (VQ-VAE) on Emulating Galaxy Images and Unsupervised Machine Learning Classification for Galaxy Morphology."

Ha, David, and Jürgen Schmidhuber. 2018. "World Models," March. https://doi.org/10.5281/zenodo.1207631.

He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. 2015. "Deep Residual Learning for Image Recognition." arXiv. http://arxiv.org/abs/1512.03385

Nelson, Kenric P. 2017. "Assessing Probabilistic Inference by Comparing the Generalized Mean of the Model and Source Probabilities." *Entropy* 19 (6). https://doi.org/10.3390/e19060286.

Nelson, Kenric P. 2024. "Photrek R3-NEW-4 Milestone Revisions".

https://docs.google.com/document/d/1rdeizBagiweonIUG6BbRZqTnp9e-CfEXn4PQoO-uZS0/edit

Nelson, Kenric P. 2024. "Risk Assessment Coupled Mean v4". Risk Assessment Coupled Mean v4.pdf

Nelson, Kenric P., Chen, Kevin, Thistleton, William J., and Clement, John. 2020. "Nonlinear-Statistical-Coupling." Photrek, LLC. https://github.com/kenricnelson/Nonlinear-Statistical-Coupling.

Zink, M., Irwin, D., Cecchet, E., Saplakoglu, H., Krieger, O., Herbordt, M., Daitzman, M., Desnoyers, P., Leeser, M., & Handagala, S. (2021). The Open Cloud Testbed (OCT): A Platform for Research into new Cloud Technologies. 2021 IEEE 10th International Conference on Cloud Networking (CloudNet), 140–147. https://doi.org/10.1109/CloudNet53349.2021.9657109