Bridging Past and Present: Reimagining Dinosaur Physiology from Fossils using 3D-to-3D Latent Diffusion

Tim Wibiral
tim.wibiral@uni-ulm.de

1 INTRODUCTION

Approximately 66 million years ago [19], an asteroid struck earth [9] and caused an environmental catastrophe that likely lead to the mass extinction at the Cretaceous/Tertiary boundary [1]. This event resulted in the extinction of all non-avian dinosaur species, creating an evolutionary niche that allowed mammals and birds to strive [6]. Today, some of these mammals attempt to reconstruct the appearance and lifestyle of dinosaurs based on the geological record. While fossilized remains provide valuable insights, certain aspects of dinosaurs' physiology are impossible to understand based on fossilized remains alone. Nevertheless, recent advancements have enabled researchers to deduce characteristics like skin color and sex for some specimens [21, 30]. Much of this progress can be attributed to the discovery of new fossils and the application of cutting-edge technologies, such as high-resolution computer tomography (CT) on feather impressions, which have allowed for the determination of feather coloring in certain cases [30].

As depicted in Figure 1, the reconstruction of extinct animals from their fossilized skeletons poses a significant challenge. This illustrates the main problem in reconstructing extinct animals from their fossilized skeletons: Almost all soft matter is lost, making a realistic reconstruction nearly impossible with existing technology.

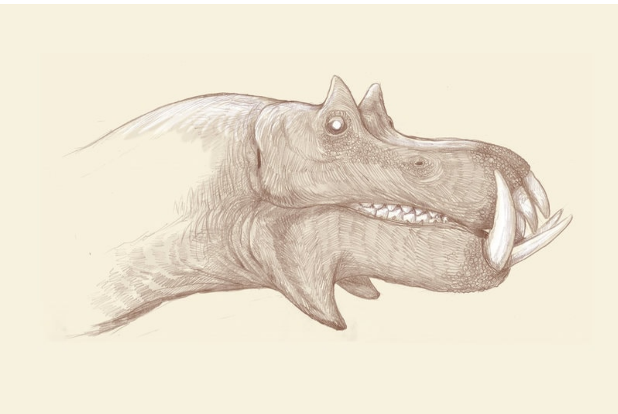

Figure 1: Speculative artistic reconstruction by C. M. Kosemen imagining how the head of a hippo would appear if reconstructed solely from its skull [7]. He is guided by the techniques used to reconstruct dinosaurs from their skeletons.

Comparing the skeletons and physiology of different living animals offers insights into their relationship and can be utilized to find out more about extinct species, for example, by identifying extended phylogenetic brackets [5]. Shared characteristics between birds and reptiles, for instance, may provide crucial clues about dinosaurs, given their evolutionary connections. It is possible that a lot of cues and unknown relationships are still hidden in the phylogenetic tree that could be exploited for more accurate representations of dinosaurs.

This leads us to the central question of our research: Can a deep learning model, trained on 3D models of living animals, effectively reconstruct the physiological appearance of dinosaurs and other extinct animals based solely on their skeletons?

To address this question, we propose a novel framework that leverages 3D-to-3D Latent Diffusion models trained on 3D scans of living animals and their skeletons to reconstruct dinosaurs' physiology based on their skeleton. In the first step, we train the latent diffusion model to generate 3D models of living animals based on the animals' skeletons. Subsequently, we apply these models to reconstruct the physiology of extinct vertebrates from their fossilized skeletons. The resulting 3D models generated by the latent diffusion model hold the potential to provide new insights into dinosaur physiology. Additional information concerning the animal, for example, the climate of its habitat[1], or external cues, like footprints, are not presented to the latent diffusion model.

2 RELATED WORK

Latent diffusion involves the application of diffusion models [22] in latent-space by auto-encoding the image-space [20]. Latent diffusion is typically used for image denoising [10] or generating new images, e.g., based on text prompts [18].

Recent advances have extended latent diffusion to generating 3D models from text prompts [14, 17, 26, 28], as well as converting 2D images to 3D models [11, 14, 23, 26]. Using pretrained text-to-image and image-to-image latent diffusion models improved performance and reduced training time [11, 17, 26, 28].

In our approach, we build upon the work of Zeng et al. [29], which applies latent diffusion to denoise 3D point clouds. We extend this technique by combining it with the pretrained text-to-3D model developed by Poole et al. [17], using cross-attention to incorporate additional information about the samples during the learning process.

3 DATASET

Due to the lack of a proper dataset, we create our own, comprising 3D models of living animals, their skeletons, and 3D models of extinct animal skeletons.[2] The dataset focuses solely on land-bound vertebrate animals, including mammals, birds, reptiles, and amphibians, to ensure comparability in size and physiology. We assume that models of fish do not improve the accuracy of the latent diffusion on land-bound animals and vice versa. All models are scaled to the same size to accommodate for the large size differences. Below, sourcing and preprocessing are described in more detail.

3.1 Sourcing the Dataset

At the moment, there is no centralized repository for 3D scans of animals, fossils, or cultural artifacts. Many museums, e.g., the

[1]Climate has a huge impact on an animal's appearance. Animals living in hot habitats like deserts often have larger extremities, while animals living in cold habitats have thicker layers of fat.

[2]The dataset is available for download at https://example.com/. Please consider the license of each model before using it. See Appendix B for more information.

Smithsonian National Museum of Natural History, create 3D catalogs of their collection, including high-resolution scans. Sometimes the data generated for papers is published online alongside the paper. While most research projects focus on small parts of a specimen, some work on scanning complete animals or fossils.[3] Appendix B contains a table with each sample used and the corresponding source and license.

3D Models of Living Animals. We acquired N^4 models from museums and M models from scientific publications. These are mostly 3D scans of sedated or dead animals, but also some were awake and hold by a human during the scanning process. We used L models of private institutions and creators. These models are often digitally constructed based on scans, images, and videos of animals. These models often depict animals in dynamic poses, such as running, but may be less accurate compared to the scans.

3D Scans and Models of Dinosaur Skeletons. Most of the around 11,000 dinosaur fossils found in the last 200 years [16] are incomplete or only consist of a few bones. If only a few bones are missing, it is possible to fill out the missing bones using other similar specimens or mirror bones from the other body half [12]. This way, it is possible to reconstruct complete models of the dinosaur skeleton.

For the 3D scans and models of dinosaur skeletons, we selected five representative specimens: Tyrannosaurus *Trix*[5] belonging to the *Naturalis Biodiversity Center* in Leiden, Netherlands, a *Dimetrodon*[6] belonging to the *Museum of Natural Sciences*, in Brussels, Belgium, a *Scleromochlus taylori*[7], a *Stegosaurus*[8] belonging to the *Science Museum of Minnesota*, in Minnesota, USA, and a *Triceratops*[9] belonging to the National Museum of Natural History, in Washington, DC, USA. Restricting the evaluation to those five dinosaur specimens allows us to conduct a thorough and high quality evaluation while still considering a diverse range of dinosaurs with distinct sizes and physiology.

For further details on the challenges and limitations of this dataset, refer to Section Section 5.

3.2 Preprocessing

To ensure uniformity in the dataset, all 3D models are normalized to fit within a cubic point cloud with a fixed length, width, and height of 2000 pixels. All models possess at least this resolution, so no model had to be upscaled. However, incorporating all models at their original size would have been impractical, given the significant differences in scale between some animals; certain specimens are more than 100 times larger than others.

To address this issue, each 3D model is proportionally scaled down to fit within the cubic point cloud using a single scaling factor, identical for all dimensions. This scaling factor is individually calculated for each model and stored for reference. During training, this scaling factor is provided to the latent diffusion model, enhancing accuracy by enabling it to learn size-related features.

While it would be possible to include additional information about the animals, such as diet, climate, or relative position in the food chain, for some living animals, this information is either not available or remains uncertain for dinosaurs. Consequently, incorporating such information in the reconstruction process could introduce errors. Nevertheless, we acknowledge that including supplementary data may be a valuable avenue for future research.

Regarding the division of 3D models of living animals, they are split into train, test, and validation set with a ratio of respectively 90%, 5%, and 5%. The 3D models of dinosaur skeletons are not included in these sets; they are used separately for evaluation by experts.

4 PROCEDURE

We adapt LION [29] for our purpose, "denoising" 3D models of skeletons to retrieve the outer shape of the animal, effectively recovering lost physiological information. The scaling factor is encoded using the pretrained text-to-3D model by Poole et al. [17] and integrated into the LION model with cross-attention, enabling utilization of the animal's size during training.

Data augmentation is randomly applied to the model input during training to enhance robustness. Gaussian noise is added to the scaling factor to account for variations in size among individuals of the same species. Additionally, a random number of bones is removed from each 3D skeleton input to simulate fossilization's corrosive effects. On average, around 2 bones are removed, with each bone having a 1% probability of being removed. Both Gaussian noise and bone removal are applied dynamically during training.

We assess the model's performance in two stages: First on the test dataset comprising 3D models of living animals, and secondly on dinosaur fossils. For the evaluation of dinosaur reconstructions, experts in the field of paleontology participate in an informal evaluation, rating each reconstructed dinosaur based on the questionnaire scales provided in Appendix A.

5 CHALLENGES AND LIMITATIONS

Working with fossils comes with many uncertainties, which present challenges and limitations for this work. Two main issues arise: the possibility of incorrect 3D skeleton models leading to erroneous predictions from the deep learning model and the potential for the deep learning model to outperform current expert reconstructions, potentially misleading assessments of its performance. The latter can largely be ruled out by comparison with well-preserved, "mummified" specimen [3, 4, 8, 24, 25]. Some of the challenges that introduce sources of error in physiological reconstructions of extinct animals are described below.

Assembling dinosaur skeletons correctly has historically been challenging, with past reconstructions exhibiting inaccuracies that have since been rectified [5]. For example, the first reconstructions of theropods in the 19th century had them walk using all four legs and placed the shoulder joints at their sides, similar to the posture of reptiles [12]. Later, the dinosaurs were put on their rear legs only, while their tale was put on the ground to support their balance [12]. Today, the posture of the legs is assembled by using computer simulation to find the most energy efficient configuration [5], resulting in theropod skeletons being

[3] e.g., The *Digital Life* project (http://digitallife3d.org/3d-model) and the exact scan of *Trix*, a T. rex Specimen (https://www.artec3d.com/cases/tyrannosaurus-rex-dinosaur-3D-scanning).

[4] *N, M,* and *L* are used as placeholders for actual numbers that cannot be provided at this point.

[5] https://sketchfab.com/3d-models/tyrannosaurus-rex-trix-8bb05194fb6d4c8a9afa3c61a58d47c7

[6] https://sketchfab.com/3d-models/dimetrodon-228350b70e7b4ec1af3987f73885efbf

[7] https://sketchfab.com/3d-models/scleromochlus-taylori-34326a87a78d4c65854537a9ca5b903f

[8] https://sketchfab.com/3d-models/stegosaurus-6cf52411f06341b3bfae6204ac23d692

[9] https://sketchfab.com/3d-models/triceratops-horridus-marsh-e9c507f179ed4455aac3b208c9e6c973

assembled with their tail not touching the ground but using it as counterbalance for their body and head [12].

While modern methods and computer simulations improve accuracy, errors can still occur. For instance, the *Field Museum of Natural History* recently updated their famous Tyrannosaurus Rex specimen *SUE* by adding a new rib cage [13].

Other challenges include the incompleteness of most fossils, deformation caused by aggregate pressure, limited information about biological sex [12, 21], substantial size differences between species, and unique features of dinosaurs not found in living animals.

6 ETHICAL CONSIDERATIONS

Generative AI training often involves unauthorized use of intellectual property and private user data [2, 27]. Concerns about this have led *Springer Nature* to prohibit the publication of AI-generated images and videos [15].

For this work, we heavily rely on models created by experts. Some 3D models used (e.g., Tyrannosaurus Rex *Trix*) are explicitly prohibited by their creators from being used for generative AI. In such cases, we sought and obtained special permission from the creators to use them for training our models. All models for which we obtained explicit permission are indicated in Appendix B with the label “Usage forbidden.” in the licensing column.

REFERENCES

[1] Luis W. Alvarez, Walter Alvarez, Frank Asaro, and Helen V. Michel. 1980. Extraterrestrial Cause for the Cretaceous-Tertiary Extinction. *Science* 208, 4448 (June 1980), 1095–1108. https://doi.org/10.1126/science.208.4448.1095 Publisher: American Association for the Advancement of Science.

[2] Gil Appel, Juliana Neelbauer, and David A. Schweidel. 2023. Generative AI Has an Intellectual Property Problem. *Harvard Business Review* (April 2023). https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem Section: Intellectual property.

[3] Victoria M. Arbour and David C. Evans. 2017. A new ankylosaurine dinosaur from the Judith River Formation of Montana, USA, based on an exceptional skeleton with soft tissue preservation. *Royal Society Open Science* 4, 5 (May 2017), 161086. https://doi.org/10.1098/rsos.161086 Publisher: Royal Society.

[4] Phil R. Bell, Federico Fanti, Philip J. Currie, and Victoria M. Arbour. 2014. A Mummified Duck-Billed Dinosaur with a Soft-Tissue Cock’s Comb. *Current Biology* 24, 1 (Jan. 2014), 70–75. https://doi.org/10.1016/j.cub.2013.11.008 Publisher: Elsevier.

[5] M. J. Benton. 2019. *Dinosaurs rediscovered: the scientific revolution in paleontology.* Thames and Hudson, New York, New York. OCLC: on1098173161.

[6] Steve Brusatte. 2022. *The rise and reign of the mammals the rise and reign of the mammals.* Picador, London, England.

[7] John Conway, C. M. Kosemen, Darren Naish, Scott Hartman, and Dan Varner. 2012. *All Yesterdays: Unique and Speculative Views of Dinosaurs and Other Prehistoric Animals.* Irregular books, s.l.

[8] Stephanie K. Drumheller, Clint A. Boyd, Becky M. S. Barnes, and Mindy L. Householder. 2022. Biostratinomic alterations of an Edmontosaurus “mummy” reveal a pathway for soft tissue preservation without invoking “exceptional conditions”. *PLOS ONE* 17, 10 (Oct. 2022), e0275240. https://doi.org/10.1371/journal.pone.0275240 Publisher: Public Library of Science.

[9] Alan R. Hildebrand, Glen T. Penfield, David A. Kring, Mark Pilkington, Antonio Camargo Z., Stein B. Jacobsen, and William V. Boynton. 1991. Chicxulub Crater: A possible Cretaceous/Tertiary boundary impact crater on the Yucatán Peninsula, Mexico. *Geology* 19, 9 (Sept. 1991), 867–871. https://doi.org/10.1130/0091-7613(1991)019<0867:CCAPCT>2.3.CO;2

[10] Jonathan Ho, Ajay Jain, and Pieter Abbeel. 2020. Denoising Diffusion Probabilistic Models. https://arxiv.org/abs/2006.11239v2

[11] Tomas Jakab, Ruining Li, Shangzhe Wu, Christian Rupprecht, and Andrea Vedaldi. 2023. Farm3D: Learning Articulated 3D Animals by Distilling 2D Diffusion. https://arxiv.org/abs/2304.10535v1

[12] Bernhard Kegel. 2018. *Ausgestorben, um zu bleiben: Dinosaurier und ihre Nachfahren* (erste auflage ed.). DuMont, Köln. OCLC: 1041912902.

[13] Field Museum. 2018. A Fresh Science Makeover for SUE - Field Museum. https://www.fieldmuseum.org/blog/fresh-science-makeover-sue

[14] Gimin Nam, Mariem Khlifi, Andrew Rodriguez, Alberto Tono, Linqi Zhou, and Paul Guerrero. 2022. 3D-LDM: Neural Implicit 3D Shape Generation with Latent Diffusion Models. https://arxiv.org/abs/2212.00842v2

[15] Nature. 2023. Why Nature will not allow the use of generative AI in images and video. *Nature* 618, 7964 (June 2023), 214–214. https://doi.org/10.1038/d41586-023-01546-4

[16] Katie Peek. 2022. Where the Dinosaurs Are. *Scientific American* 324 (2022), 72. https://doi.org/10.1038/scientificamerican0221-72

[17] Ben Poole, Ajay Jain, Jonathan T. Barron, and Ben Mildenhall. 2022. DreamFusion: Text-to-3D using 2D Diffusion. https://arxiv.org/abs/2209.14988v1

[18] Aditya Ramesh, Mikhail Pavlov, Gabriel Goh, Scott Gray, Chelsea Voss, Alec Radford, Mark Chen, and Ilya Sutskever. 2021. Zero-Shot Text-to-Image Generation. https://doi.org/10.48550/arXiv.2102.12092 arXiv:2102.12092 [cs].

[19] Paul R. Renne, Alan L. Deino, Frederik J. Hilgen, Klaudia F. Kuiper, Darren F. Mark, William S. Mitchell, Leah E. Morgan, Roland Mundil, and Jan Smit. 2013. Time Scales of Critical Events Around the Cretaceous-Paleogene Boundary. *Science* 339, 6120 (Feb. 2013), 684–687. https://doi.org/10.1126/science.1230492 Publisher: American Association for the Advancement of Science.

[20] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. 2021. High-Resolution Image Synthesis with Latent Diffusion Models. https://arxiv.org/abs/2112.10752v2

[21] Mary Higby Schweitzer, Craig Johnson, Thomas G. Zocco, John R. Horner, and Jean R. Starkey. 1997. Preservation of biomolecules in cancellous bone of Tyrannosaurus rex. *Journal of Vertebrate Paleontology* 17, 2 (June 1997), 349–359. https://doi.org/10.1080/02724634.1997.10010979 Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/02724634.1997.10010979.

[22] Jascha Sohl-Dickstein, Eric Weiss, Niru Maheswaranathan, and Surya Ganguli. 2015. Deep Unsupervised Learning using Nonequilibrium Thermodynamics. In *Proceedings of the 32nd International Conference on Machine Learning*. PMLR, 2256–2265. https://proceedings.mlr.press/v37/sohl-dickstein15.html ISSN: 1938-7228.

[23] Gabriela Ben Melech Stan, Diana Wofk, Scottie Fox, Alex Redden, Will Saxton, Jean Yu, Estelle Aflalo, Shao-Yen Tseng, Fabio Nonato, Matthias Muller, and Vasudev Lal. 2023. LDM3D: Latent Diffusion Model for 3D. https://arxiv.org/abs/2305.10853v2

[24] Charles H. Sternberg. 1908. Expedition to the Laramie Beds of Converse County, Wyoming. *Transactions of the Kansas Academy of Science (1903-)* 22 (1908), 113–116. https://doi.org/10.2307/3624729 Publisher: Kansas Academy of Science.

[25] Charles H. Sternberg. 1909. A New Trachodon from the Laramie Beds of Converse County. *Science* 29, 749 (1909), 753–754. https://doi.org/10.1126%2Fscience.29.749.747

[26] Jiaxiang Tang. 2022. Stable-dreamfusion: Text-to-3D with stable-diffusion.

[27] James Vincent. 2022. The scary truth about AI copyright is nobody knows what will happen next. https://www.theverge.com/23444685/generative-ai-copyright-infringement-legal-fair-use-training-data

[28] Haochen Wang, Xiaodan Du, Jiahao Li, Raymond A. Yeh, and Greg Shakhnarovich. 2023. Score Jacobian Chaining: Lifting Pretrained 2D Diffusion Models for 3D Generation. (2023), 12619–12629. https://openaccess.thecvf.com/content/CVPR2023/html/Wang_Score_Jacobian_Chaining_Lifting_Pretrained_2D_Diffusion_Models_for_3D_CVPR_2023_paper.html

[29] Xiaohui Zeng, Arash Vahdat, Francis Williams, Zan Gojcic, Or Litany, Sanja Fidler, and Karsten Kreis. 2022. LION: Latent Point Diffusion Models for 3D Shape Generation. *Advances in Neural Information Processing Systems* 35 (Dec. 2022), 10021–10039. https://proceedings.neurips.cc/paper_files/paper/2022/hash/40e56dabe12095a5fc44a6e4c3835948-Abstract-Conference.html

[30] Fucheng Zhang, Stuart L. Kearns, Patrick J. Orr, Michael J. Benton, Zhonghe Zhou, Diane Johnson, Xing Xu, and Xiaolin Wang. 2010. Fossilized melanosomes and the colour of Cretaceous dinosaurs and birds. *Nature* 463, 7284 (Feb. 2010), 1075–1078. https://doi.org/10.1038/nature08740 Number: 7284 Publisher: Nature Publishing Group.

A QUESTIONNAIRES FOR RATING THE DINOSAUR RECONSTRUCTIONS

Here, we would describe how experts could rate the reconstructed dinosaurs. The questionnaire would contain different scales to rate different aspects of the physiology, like possession of lips, distribution of fat and muscles, or the general proportions.

B MODEL SOURCES

We list all data sources and the models associated with them. “Usage forbidden.” means that the models are explicitly excluded from usage in datasets for Generative AI. If we realized this proposal, we would have to ask for explicit permission for those models.

The three tables list all 3D models used. Table 1 contains the 3D scans of fossilized dinosaur skeletons used. Table 2 and Table 3 list the 3D scans of animals shapes and skeletons respectively.

Species	License	URL
Dimetrodon	Usage forbidden.	https://sketchfab.com/3d-models/dimetrodon-228350b70e7b4ec1af3987f73885efbf
Scleromochlus	Not provided.	https://sketchfab.com/3d-models/scleromochlus-taylori-34326a87a78d4c65854537a9ca5b903f
Stegosaurus	Not provided.	https://sketchfab.com/3d-models/stegosaurus-6cf52411f06341b3bfae6204ac23d692
Triceratops	Public domain.	https://sketchfab.com/3d-models/triceratops-horridus-marsh-e9c507f179ed4455aac3b208c9e6c973
Tyrannosaurus (“Trix”)	Usage forbidden.	https://sketchfab.com/3d-models/tyrannosaurus-rex-trix-8bb05194fb6d4c8a9afa3c61a58d47c7

Table 1: 3D models based on scans of dinosaur fossils.

Species	License	URL
Black-Legged Poison Frog	CC BY-NC	https://sketchfab.com/3d-models/model-19-black-legged-poison-frog-ab1f6c3c5229421d83b3e68443d566c9
Southern White Rhino	CC BY-NC	https://sketchfab.com/3d-models/model-56a-southern-white-rhino-8e97b62a90f44ce19ea9e3fd421f55b4
Sling Tailed Agama	CC BY-NC	https://sketchfab.com/3d-models/model-59a-sling-tailed-agama-22f56f22ae564cdc8ef0242c7b08456b
⋮	⋮	⋮

Table 2: 3D models of non-extinct animals.

Species	License	URL
Emporer Penguin	Usage forbidden.	https://sketchfab.com/3d-models/emperor-penguin-imnh-r-1367-62352f57731c48b29320643fba3e5dda
Horse	CC Attribution	https://sketchfab.com/3d-models/horse-skeleton-eaca504567604e879b8ab2cf2763025
Horned Marsupial Frog	Not provided.	https://sketchfab.com/3d-models/model-56a-southern-white-rhino-8e97b62a90f44ce19ea9e3fd421f55b4
⋮	⋮	⋮

Table 3: 3D models of non-extinct animals’ skeletons.