Manuscript for

# Challenges and opportunities in applying AI to evolutionary morphology

Yichen He (1)*, James M. Mulqueeney (1, 2)*, Emily C. Watt (1, 3)*, Arianna Salili-James (4, 5), Nicole S. Barber (1, 6), Marco Camaiti (1), Eloise S. E. Hunt (1, 7, 8), Oliver Kippax-Chui (1, 8, 9), Andrew Knapp (1, 10), Agnese Lanzetti (1, 11), Gizéh Rangel-de Lázaro (1, 12, 13), Jack K. McMinn (1, 14), Joseph Minus (1, 15), Ashwini V. Mohan (1, 16), Lucy E. Roberts (1), Darcy Adhami (1, 7, 17), Enrico Grisan (18), Qianqian Gu (4, 5), Victoria Herridge (1, 19), Sanson T. S. Poon (4, 5), Taylor West (10, 17), Anjali Goswami (1)

* Joint first-authors

Corresponding author: Yichen He (yichen.he@nhm.ac.uk)

(1) Life Sciences, Natural History Museum, London, UK

(2) Department of Ocean & Earth Science, National Oceanography Centre Southampton, University of Southampton, UK

(3) Division of Biosciences, University College London, UK

(4) AI and Innovation, Natural History Museum, London, UK

(5) Digital, Data and Informatics, Natural History Museum, London, UK

(6) Department of Anthropology, University College London, UK

(7) Department of Life Sciences, Imperial College London, London, UK

(8) Grantham Institute, Imperial College London, London, UK

(9) Department of Earth Science and Engineering, Imperial College London, London, UK

(10) Centre for Integrative Anatomy, University College London, UK

(11) School of Geography, Earth and Environmental Sciences, University of Birmingham, UK

(12) School of Museum Studies, University of Leicester, UK

(13) School of Oriental and African Studies, SOAS, UK

(14) Department of Earth Sciences, University of Oxford, UK

(15) School of Biological and Behavioural Sciences, Queen Mary University of London, UK

(16) Biodiversity Genomics Laboratory, Institute of Biology, University of Neuchâtel, Switzerland

(17) Imaging and Analysis Centre, Natural History Museum, UK

(18) School of Engineering, London South Bank University, UK

(19) School of Biosciences, University of Sheffield, UK

ORCiDs
Yichen: https://orcid.org/0000-0003-3464-7526
James: https://orcid.org/0000-0003-3502-745X
Emily: https://orcid.org/0000-0003-0214-7105
Arianna: https://orcid.org/0000-0003-1125-2054
Nicole: https://orcid.org/0000-0003-3517-9672
Marco: https://orcid.org/0000-0002-7388-4586
Eloise: https://orcid.org/0000-0002-3731-3479
Oliver: https://orcid.org/0000-0003-4551-6281
Andy: https://orcid.org/0000-0003-4822-5623
Agnese: https://orcid.org/0000-0003-1364-0544
Gizeh: https://orcid.org/0000-0002-3257-0783
Jack: https://orcid.org/0000-0003-2696-9306
Joseph: https://orcid.org/0009-0003-3968-4382
Lucy: https://orcid.org/0000-0002-1293-8667
Ashwini: https://orcid.org/0000-0001-9505-0665
Darcy: n/a
Enrico: https://orcid.org/0000-0002-7365-5652
Qianqian: https://orcid.org/0000-0002-9329-6554
Victoria: https://orcid.org/0000-0001-5204-8426
Sanson: https://orcid.org/0000-0002-7158-5400
Taylor: https://orcid.org/0000-0001-8933-3773
Anjali: https://orcid.org/0000-0001-9465-810X

# Main Title

Challenges and opportunities in applying AI to evolutionary morphology

# Running Title

AI for evolutionary morphology

# Abstract

Artificial intelligence (AI) is poised to transform many aspects of society, and the study of evolutionary morphology is no exception. Machine learning-grade methods of AI such as Principal Component Analysis (PCA) and Cluster Analysis have been commonplace in evolutionary morphology for decades, but the last decade has seen increasing application of Deep Learning to ecology and evolutionary biology, opening up the potential to circumvent longstanding barriers to rapid, big data analysis of phenotype. Here we review the current state of AI methods available for the study of evolutionary morphology and discuss the prospectus for near-term advances in specific subfields of this research area, including the potential of new AI methods that have not yet been applied to the study of morphological evolution. We introduce the main available AI techniques, categorising them into three stages based on their order of appearance: (i) Machine Learning, (ii) Deep Learning with neural networks and (iii) the most recent advancements in large-scale models and multimodal learning. Next, we present existing AI approaches and case studies using AI for evolutionary morphology, including image capture and segmentation, feature recognition, morphometrics, phylogenetics, and biomechanics. Finally, we discuss areas where there is potential, but no current application of AI to key areas in evolutionary morphology. Combined, these advancements and potential developments have the capacity to transform the evolutionary analysis of organismal phenotype into evolutionary phenomics, launch it fully in the “Big Data" sphere, and align it with genomics and other areas of bioinformatics.

**Keywords**

Artificial Intelligence

Phenomics

Evolutionary Morphology

Machine Learning

# Introduction

The rapid proliferation of tools using artificial intelligence (AI) has highlighted both its immense potential and the numerous challenges its implementation faces in the biological sciences. Traditional AI methods (i.e. Machine Learning) have been widely used in biology for decades; indeed, common analytical methods such as Principal Component Analysis (PCA) and Cluster Analysis are both types of Machine Learning (ML). Deep Learning (DL) has gained significant traction since the early 2010s and is increasingly applied to biological problems, including image analysis (Akçakaya et al., 2022; Angermueller et al., 2016; Hallou et al., 2021; Li et al., 2023; Zhichao Liu et al., 2021; Moen et al., 2019; Pratapa et al., 2021; Ravindran, 2022; "What's next for bioimage analysis?," 2023), molecular analysis (Atz et al., 2021; Audagnotto et al., 2022; Korfmann et al., 2023; Kuhn et al., 2021; Kwon et al., 2021), and a broad range of topics in ecology and evolutionary biology (Borowiec et al., 2022; Lürig et al., 2021; Pichler and Hartig, 2023).

One key area overlapping with many topics in evolutionary biology, is the field of evolutionary morphology, which aims to characterise and reconstruct the evolution of organismal phenotypes. The scope of evolutionary morphology is huge, encompassing pattern, process, and mechanism, from cellular to macroevolutionary levels, across the entire 3.7-billion-year history of life on Earth and, consequently, often involves massive datasets. Due to the sheer quantity of potentially informative data, some of the most significant challenges evolutionary morphologists face is the ability to collect, process and analyse this data in a reasonable time frame whilst limiting computational cost. Researchers often face a trade-off between the breadth and depth of their study, as, typically, high-resolution morphological datasets must sacrifice taxonomic, ecological or chronological coverage owing to time- and computational limitations. AI offers an unparalleled opportunity to bridge this breadth-depth gap and thus transform the field into "Big Data" science, thereby supporting the development of *evolutionary phenomics*. By making big data analysis more feasible, integrating AI into this field will ultimately allow a better understanding of the drivers and mechanisms of morphological evolution.

Here, we focus on the applications of AI to the study of evolutionary morphology, exploring not only existing uses but also the potential of recent AI methods that have not yet been applied to the study of morphological evolution. We introduce the main available AI techniques, categorising them into three groups based on their order of appearance: (i) Machine Learning, (ii) Deep Learning with Neural Networks, and (iii) Recent advancements from Transformers to large-scale models. Next, we present existing AI approaches in the order of a common lifecycle of evolutionary morphological studies: (i) Data Acquisition, (ii) Image Data Processing, (iii) Phenomics, (iv) Evolutionary applications. We also focus on six case studies in which AI can benefit evolutionary morphological studies. Finally, we discuss areas where there is potential but no current application of AI to key areas in evolutionary morphology.

# Evolution of AI methods

Here, we outline the major stages in the evolution of AI relevant to the study of evolutionary morphology. We categorise them into three imperfect groupings based on their order of appearance: (i) Machine Learning, (ii) Deep Learning with Neural Networks, and (iii) Recent advancements from Transformers to large-scale models.

We begin by providing the key definitions necessary for a base level understanding of this review. These primarily centre on the nested relationships of AI, ML, and DL (Figure 1), but also include the adjacent and overlapping field of computer vision. Because AI applications for evolutionary morphology primarily involve the analysis of images or text, computer vision is often an integral part of AI applications to evolutionary morphology, including most of those discussed here. However, it is worth noting that computer vision is not limited to AI but also present in numerous applications for image data that do not involve AI (Samoili et al., 2020). Further methodological definitions are provided where required in the main text.

***Artificial intelligence***, or AI, concerns intelligent machines capable of mimicking human-like cognitive functions. It is, however, particularly challenging to specifically define AI as its scope is extremely broad. The European AI strategy (European Commission, 2018) provides a definition as follows: "Artificial Intelligence refers to systems that display intelligent behaviour by analysing their environment and taking action — with some degree of autonomy — to achieve specific goals", leaving the interpretation of *intelligent behaviour* open to the reader. Russel and Norvig (2021) try to provide a more operative definition of AI, as a system that can either 'reason' or act human-like, or reason or act rationally.

***Machine learning***, or ML, is a subset of AI, and can be defined as "the ability of systems to automatically learn, decide, predict, adapt, and react to changes, improving from experience and data, without being explicitly programmed" (Amalfitano et al., 2024).

***Deep learning***, or DL, is, in turn, a subset of ML wherein learning is achieved through complex neural networks designed to simulate the cognitive architecture of the brain. Fine-grained tasks on complex data can be achieved using vast amounts of data and with limited human intervention.

***Computer vision*** is a multidisciplinary field of computer science that enables machines to interpret, analyse, and understand visual information from the world, mimicking human vision capabilities through image and video processing algorithms. It refers to using computers for object-class recognition, where objects or individuals can be identified in 2D and 3D digital media. While many applications of computer vision for evolutionary morphology involve AI, it is not limited to AI and is applied in diverse fields.

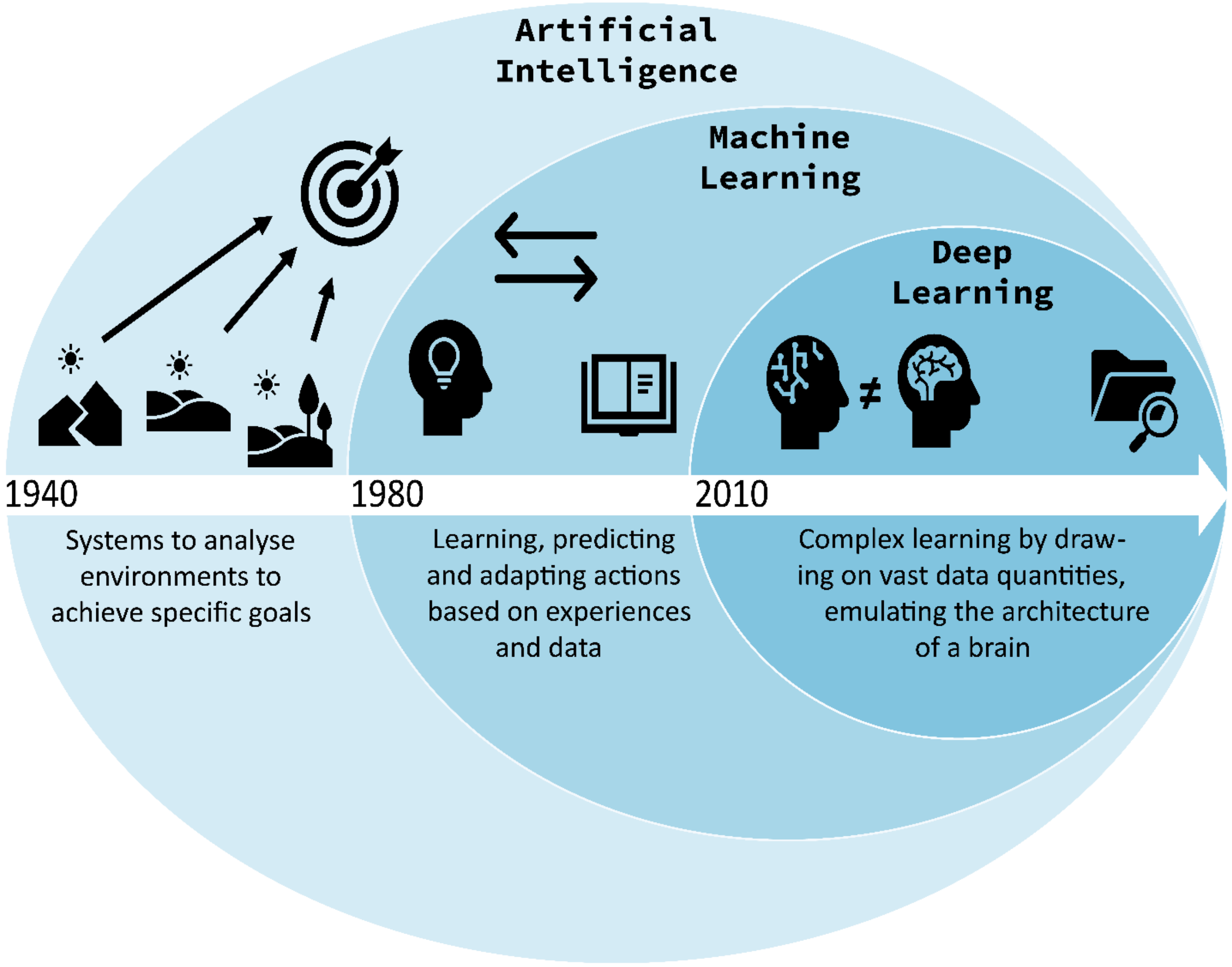

Figure 1: Broad definitions, relationships, and differences between artificial intelligence (AI), Machine Learning (ML) and Deep Learning (DL), the sequential development of each successive subset, and their broad introductions over time (Carbonell et al., 1983; Goodfellow et al., 2016).

## Machine Learning in Computer Vision

Classical (prior to DL) computer vision pipelines were composed of two separate computational steps. The first involved the extraction of local or global characteristics (features) that were deemed useful for a task from images. This meant that, for example, the borders and edges of an image needed to be identified, and subsequently, an object could be detected based on the edges, as in the active contours (Kass et al., 1988) and level sets methods (Chan and Vese, 1999; Osher and Sethian, 1988). The extracted features were then used as inputs to ML algorithms that were optimised for structured data, e.g. tabular data.

Subsequent efforts were then devoted to the design of methods to extract relevant features, i.e. features that were able to capture the relevant structures within an image, such as Haar features (Papageorgiou et al., 1998), Scale-Invariant Feature Transform (SIFT) (Lowe, 2004), Histogram of Oriented Gradients (HOG) (Dalal and Triggs, 2005) Fisher kernels (Perronnin et al., 2010; Perronnin and Dance, 2007), and curvelets (Candès et al., 2006). These engineered (or hand-crafted, or heuristic) features were then often used as inputs for ML methods, which can be broadly classified into the following approaches: predictive methods, classification, clustering, and dimension reduction (Breiman, 2001; Cortes and Vapnik, 1995; Jolliffe and Cadima, 2016; Lloyd, 1982). Although DL architectures and convolutional neural networks (CNNs) had already been proposed in the early 1990s (LeCun et al., 1989), their success was limited due to a lack of computational power and the availability of large datasets needed to fully exploit their capabilities. However, there were some attempts to design ML systems that could learn the extraction of optimal linear features for the downstream task (classification, detection, clustering, reduction) within a boosting framework (Vedaldi et al., 2007).

## Deep Learning

Although artificial neurons (McCulloch and Pitts, 1943) and then artificial neural networks were introduced several decades ago (Rosenblatt, 1958), they were often outperformed by other methods, especially ensembles of decision trees like Random Forests (Breiman, 2001) or boosted trees (Chen and Guestrin, 2016) across a variety of tasks. This failure was mainly due to the difficulty in training fully connected networks (networks in which the neurons of each layer are connected to all neurons in the following layer) with more than few layers. Even when shared-weights approaches and CNNs were introduced (Fukushima, 1980; LeCun et al., 1989), they remained on the fringe of the computer vision community, with the primary bottlenecks being the computational power required to build networks with multiple layers and the amount of data needed to train such systems.

As the availability of data and the performance of computer hardware improved, especially with the advent of graphics processing units (GPUs), deep CNNs rose to prominence in the field of computer vision. The year 2012 represents a key turning point, when a deep CNN achieved the best result in the ImageNet Large Scale Visual Recognition Challenge (classifying millions of images into 1000 classes) (Krizhevsky et al., 2012). Ever since, computer vision tasks have been dominated by solutions using deep artificial neural networks, to the extent that learning with deep neural networks (DL) is now generally referred to as AI, a name formerly used only for methods trying to solve general intelligence tasks, rather than specific tasks. In recent years, DL has undergone significant expansion into diverse domains, demonstrating its adaptability and offering promising solutions to challenges in various fields such as physics, medicine, and even gaming (Poon et al., 2023; Raissi et al., 2019; Shallue and Vanderburg, 2018; Silver et al., 2016).

Around the same time, neural network-based methods such as recurrent neural network (RNN) (Graves et al., 2013) and Long-Short-Term-Memory (LSTM) (Hochreiter and Schmidhuber, 1996) have been applied to sequential data, and have shown great results for handling text and time series data. These methods have then been widely used in natural language processing (NLP) tasks.

The difficulty of gathering a big enough dataset to fully train a DL model for a specific task can be mitigated by the assumption that many low-level features learned by large models are generally enough for most tasks. Under this assumption, the features learnt for a task can also be used (transferred) in a different task. A technique frequently used in DL is the use of pre-trained models that are then fine-tuned (the entire model adapts to the new task) or used for transfer learning (only the final layers of the models are trained). Using pre-trained models reduces the need for large datasets, often improves model performance, and saves training time and resources. A common example is the use of models pre-trained with the ImageNet dataset for downstream tasks (Chen et al., 2017; Ren et al., 2016), such as in Sun *et al*. (2018), where the ImageNet-based model was used for object detection from underwater videos in marine ecology.

## Transformer, large-scale AI models, and Multimodal Learning

In 2017, a model architecture known as the Transformer was developed to address many NLP tasks, such as translation (Vaswani et al., 2017; Vydana et al., 2021). Transformer uses a self-attention mechanism, allowing each token (i.e., words, phrases, sentences, etc.) to interact with other tokens during training. Transformer can handle more information than RNNs and LSTM, can analyse contextual information, and is also better at parallelisation. Since Transformer's introduction, it has become the state-of-the-art for many NLP tasks (Ahmed et al., 2017; Baevski and Auli, 2019).

By 2020, most vision models were using CNN-based methods. Transformer has started being implemented as the backbone architecture for vision models (Dosovitskiy et al., 2021; Ze Liu et al., 2021). A common method is to divide an image into patches, which are treated as sequential inputs similar to tokens in NLP tasks. When Transformer is applied, models can recognise patterns and relationships between different parts of the image.

Research has shown that having large and diverse datasets allows models to generalise well and perform more accurately (Goodfellow et al., 2016; Russakovsky et al., 2015). Supervised learning is a common learning strategy that requires all training data to be manually labelled. However, gathering a large amount of labelled data is extremely labour-intensive. Different training strategies are applied to tackle this problem (Figure 2). Semi-supervised learning uses both labelled and unlabelled data for training (Zhu and Goldberg, 2022). Weakly-supervised learning uses less accurately labelled data for training (Lin et al., 2016). Self-supervised learning only uses unlabelled data. These strategies allow DL models to leverage as much data as possible without the need for extensive manual work.

Self-supervised learning has been widely used in NLP studies. One example uses parts of sentences as input data to predict entire sentences, thereby allowing all the unlabelled text to be considered as training data (Devlin et al., 2019). Models trained with masked sentences can be used as powerful pre-trained models for fine-tuning downstream tasks. With access to more training data and larger model architectures, generative models like the Generative Pre-trained Transformer (GPT) family were developed (Brown et al., 2020; Radford et al., 2019, 2018). Recent GPT models (e.g., GPT-3.5 and GPT-4) are capable of performing exceptionally well on many NLP tasks, even when doing zero-shot (no training needed for new tasks) or few-shot (only a few training samples needed) learning.

Contrastive learning is one of the self-supervised learning strategies that is widely used in computer vision (Oord et al., 2019; Wu et al., 2018). The idea of contrastive learning is to train a model to map similar instances (e.g., a different view of the same image) close together, while mapping dissimilar images farther apart in the feature space. Although different ways have been designed to map similar/dissimilar instances (Chen et al., 2020; He et al., 2020), the fundamental concept remains the same. As a result, contrastive learning enables models to capture intricate visual patterns and semantics from data without the need for labelled data, thereby improving performance on downstream tasks. Later, masked images (where parts of images are obscured) have been used to predict original images and have been shown to achieve promising results (He et al., 2021).

These learning strategies have opened new avenues for training models, which enable the training of large models using unlabelled or a small set of labelled data, which is particularly applicable to biological sciences given the wealth of data available in natural history collections. Additionally, AI has been successfully applied to process various data modalities, including text, images, and videos. Multimodal learning can be implemented by combining features extracted from different data modalities into one feature space. Multimodal learning enables tasks such as generating images with text descriptions or generating descriptions for images (Radford et al., 2021). With more data available (e.g., through self-supervised learning) and the advancement of AI models (e.g., Transformer), the field of multimodal learning is rapidly evolving. In evolutionary morphology, multimodal learning can effectively process diverse data modalities, such as photographs, micro-computed tomography (micro-CT) scans, and 3D mesh models (Figure 3).

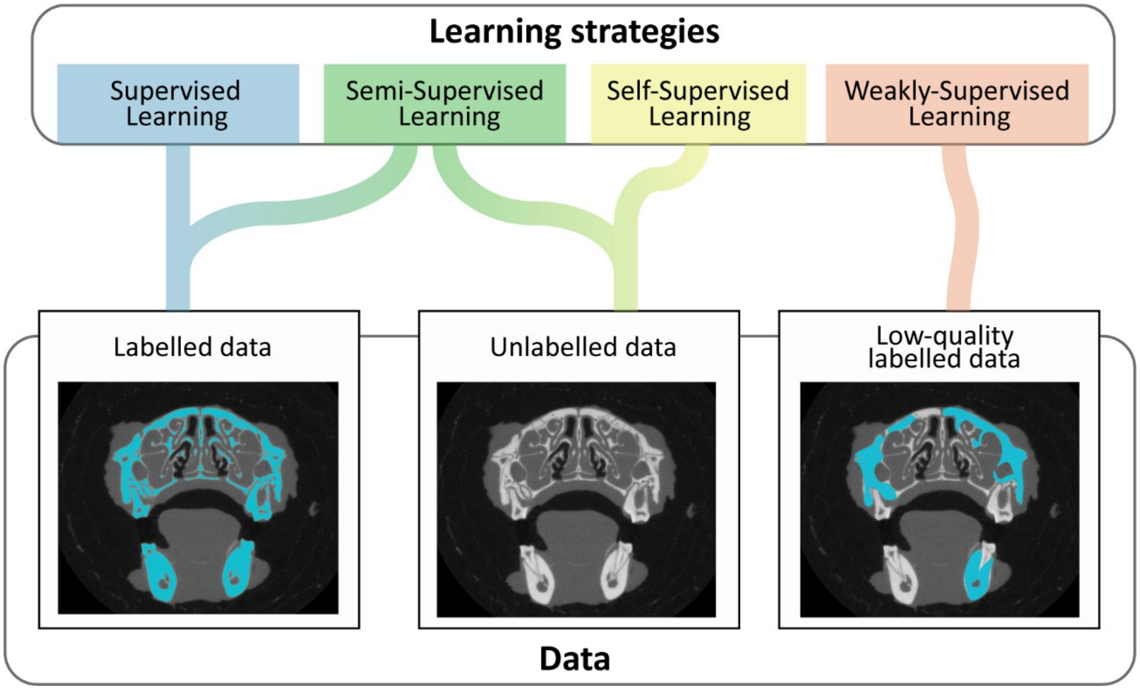

Figure 2: An overview of existing learning strategies and the levels of labelling used in these strategies.

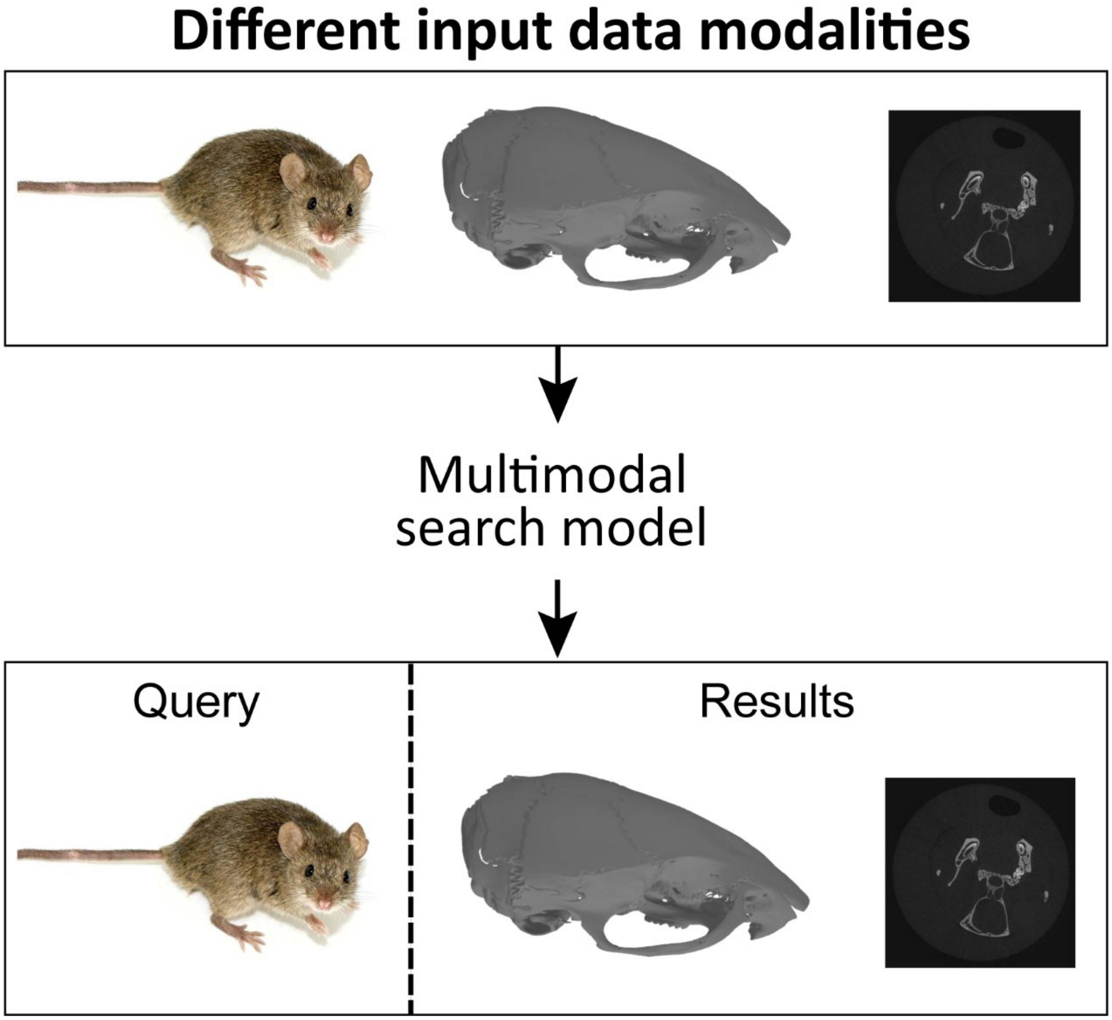

Figure 3: An example of a multimodal search. By inputting a photograph, a multimodal search model will find the mesh models and Micro-CT scans that share similar features with the input photograph.

Here, we have only briefly reviewed three major stages in the development of AI. A full review is beyond the scope of this paper, and there are numerous other subfields of AI not explicitly reviewed in this section, such as robotics and graph neural networks. Nonetheless, these methods hold substantial potential for the study of evolutionary morphology and, where appropriate, will be noted in the subsequent sections focused on current usage and future applications in this field.

# AI for Evolutionary Morphology

In this next section, we pivot towards a goal-oriented review and prospectus of applications of AI in evolutionary morphology, with accompanying case studies. We present the overview of currently available AI tools for evolutionary morphology studies in four sections: Data Acquisition, Image Data Processing, Phenomics, and Applications to Research in Evolutionary Morphology. We introduce these methods with a schematic of generalised AI workflows (Figure 4) which are expanded up in the sections that follow.

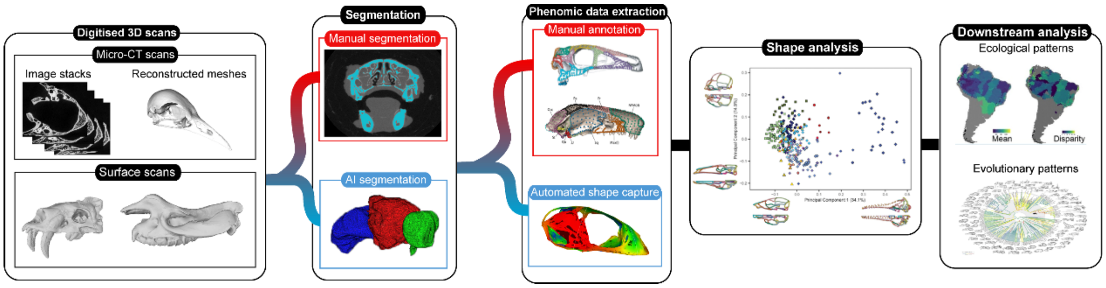

Figure 4: Schematic of a common workflow using manual and AI approaches for evolutionary morphological analysis involving 3D images. Manually annotated caecilian and theropod skulls from Bardua et al. (2019) and Felice et al. (2020), Shape information and evolutionary patterns figures from Goswami et al. (2022). Ecological analysis figure from Foister and Felice (2021)

## Data Acquisition

The first step of acquiring data is to collect the relevant samples which are to be used in the subsequent investigation in an appropriate and ethical way. For analysis of evolutionary morphology this includes not only obtaining the data that is being measured but also the corresponding metadata such as details about museum collections. The suitability, quality, and quantity of data are of critical importance to the development and implementation of AI models. Data should be diverse and clean; fulfilling these requirements can make a larger difference than model choice, and without data that conform to these requirements, good models will perform badly (Whang et al., 2023). The diversity of data refers to including enough examples of each class of interest. Cleaning data is the process of minimising error from training datasets. Preprocessing a dataset increases the suitability of the data for training and can include contrast enhancement, noise reduction and masking, where a portion of the image is designated for further analysis (Lürig et al., 2021). Determining how much data is enough depends on the specific problem at hand. Scarce data can be expanded using existing databases or by employing pre-trained networks for transfer learning (Sharif Razavian et al., 2014). However, DL models can be successful on small training sets. Few-shot learning is a form of transfer learning that uses training data where 1-20 examples of each class are available (Y. Wang et al., 2021). Scarce data for a small number of classes is commonly referred to as the long-tail problem. Where there is imbalance between the presence of classes in the dataset, the model may find it difficult to discriminate the scarcely represented classes and perform unreliably (Schneider et al., 2020).

Data scarcity and imbalance can be improved by additional data collection, or artificial data expansion, e.g. augmentation. Alternatively, imbalance can be tackled by explicitly accounting for biases in the training algorithm (Buda et al., 2018). Augmentation effectively increases the size of the training set without new data collection, by distorting images to create ‘new’ images from the existing data. This can be achieved by rotating, mirroring, scaling or by altering the pixel values (Shorten and Khoshgoftaar, 2019). This process must be controlled with the aim of the model in mind. For example, for planktonic foraminifera, the chirality of a species can be important in species classification, meaning augmentation by mirroring distorts the labelled image into a facsimile of a different species (Hsiang et al., 2019).

## Identifying and cataloguing specimen data

Many, perhaps even most, studies of evolutionary morphology are based primarily on data housed within museum collections. However, museum collections are rarely fully catalogued and even then it is difficult to search for a specific specimen or representatives of specific groups. This difficulty is because data is often inconsistent in quality and structure, particularly in large collections (Dutia and Stack, 2021). AI can play a key role in this, particularly when it comes to tasks of identifying, cataloguing, and locating specimens within collections. Some of the key challenges to solve within that topic include recognising species and extracting taxonomic and metadata to enable effective searches.

DL has recently been applied to many types of biological specimens and collections (e.g. Soltis et al., 2020). These methods have been developed and applied extensively for recognising species, metadata, traits, and even life history stages of digitised specimens (Case Study 1). This is most established in the botanical sciences, where flat herbarium sheets are easily digitised in large numbers, likely due to their relative ease of digitisation. In some instances, albeit to a lesser degree, species identification methods have also been applied to digitised photographs of animal collections  (e.g., Ling et al 2023; Macleod 2017). Applications to species identification of both plants and animals from photographs have been greatly enhanced by citizen science, resulting in useful online tools such as iNaturalist and Pl@ntNet (Goëau et al., 2013; Unger et al., 2021). CNN algorithms have borne promising results and can correctly distinguish morphologically similar species (Feng et al., 2021; Hollister et al., 2023). Other machine-learning methods, such as those described by Wilson et al. (2023), have also been applied to rescaling and increasing the quality of, and extracting metadata from, images of museum specimens, allowing for automatic feeding of this information into databases.

Beyond images of the specimens themselves, AI approaches to capturing information of specimen labels is a critical aspect of cataloguing specimens and making key data searchable. Case Study 1b outlines the approach by which ML can be used to identify labels and transcribe them for databases, saving vast amounts of manual effort. Together, species identification and taxonomic and meta-data extraction methods from images represent a powerful tool for unlocking the full potential of natural history collections. These approaches can make data more discoverable and usable for documenting biodiversity both in collections and in the field (Karnani et al., 2022; Schuettpelz et al., 2017; Wäldchen and Mäder, 2018; White et al., 2020).

Information on specimens is not limited to museum catalogues, but is also available in the wealth of scientific publications detailing and imaging specimens for varied purposes. However, extracting taxonomic data from the literature to describe or identify living and fossil species is a time-consuming task. Often it is also difficult to find the first appearance of a species name and correctly identify all synonyms for a taxon, as well as accounting for more recent taxonomic reclassifications. Recently, a few research groups have attempted to tackle this problem using ML, with both NLP and other Deep Neural Network (DNN) algorithms having been successfully applied to extract scientific terms and taxonomic names from scientific articles. This is a relatively new application of ML and more work is required to train models on a variety of sources, including articles in different languages and historic publications (Le Guillarme and Thuiller, 2022).

Once these data are captured, we need effective tools for searching for connected specimens. ML has not yet been adopted on a large enough scale to allow searching global natural history collections and connecting specimens. Dutia and Stack (2021) recommend ‘Heritage Connector’, a framework and software for using ML to allow better connecting specimens in collections and publications. This software achieved a precision score of greater than 85% with science museum group records. If refined or applied on a wider scale to natural history collections, it will certainly ease access to the vast specimen data available in global collections.

---

## Case Study 1: Machine Learning within Museum Digitisation and data collection

The digitisation of museum specimens is vital for the future of collections, and their datasets undoubtedly play a significant role in scientific research in many fields, such as evolutionary morphology. In recent years, institutions have increasingly begun to incorporate recent technologies and recent ML tools within their digitisation pipelines. These implementations have led to a range of advances from speeding up digitisation processes, enabling the digitisation of items that were once difficult to digitise, to unlocking novel data from post-processing digitised items.

**Example 1a: Machine Learning & Robotics for Specimen Digitisation**

At its most basic definition, digitisation involves the creation of digital objects from physical items, and, within museums, this is often attributed to the photographing, scanning, or filming of physical specimens. However, traditional ways of digitising artefacts, such as digitising each specimen individually, can undoubtedly be invasive to the specimen, highly time-consuming and not very cost effective. This has led to a series of innovations that can help advance museum digitisation, from drawer scanning (Schmidt et al., 2012), which enables multiple specimens to be digitised at once, to special rotating platforms that, when combined with photogrammetry techniques allow for the 3D scanning of specimens, whilst avoiding the use of more expensive or time-consuming scanning techniques (Medina et al., 2020). ML can lend a hand to these innovations to advance digitisation even more, such as the use of computer vision techniques and CNNs to segment individual specimens from whole-drawer scans (Blagoderov et al., 2012; Hansen et al., 2020; Hudson et al., 2015).

Another technological advancement that can aid digitisation is robotics. Robots are indeed already in use in other sectors such as book scanning at libraries (Dumiak, 2008). Though usually highly expensive, prices of robotic arms have been decreasing, and one can now purchase a robotic arm for less than £20,000 (Stanford University, 2022). This has enabled digitisation teams within museums such as the Natural History Museum, London, to start exploring robotics for digitisation research (Scott et al., 2023). Here, the goal is to have a collaborative robot (cobot) aid a digitiser in the mass digitisation of certain specimens (Figure 5). By implementing CNN algorithms and/or turning to reinforcement learning (RL), a robotic arm can lead to a pipeline that can enable digitisation teams to mass digitise multitudes of specimens, even possibly overnight, hence revolutionising museum digitisation work.

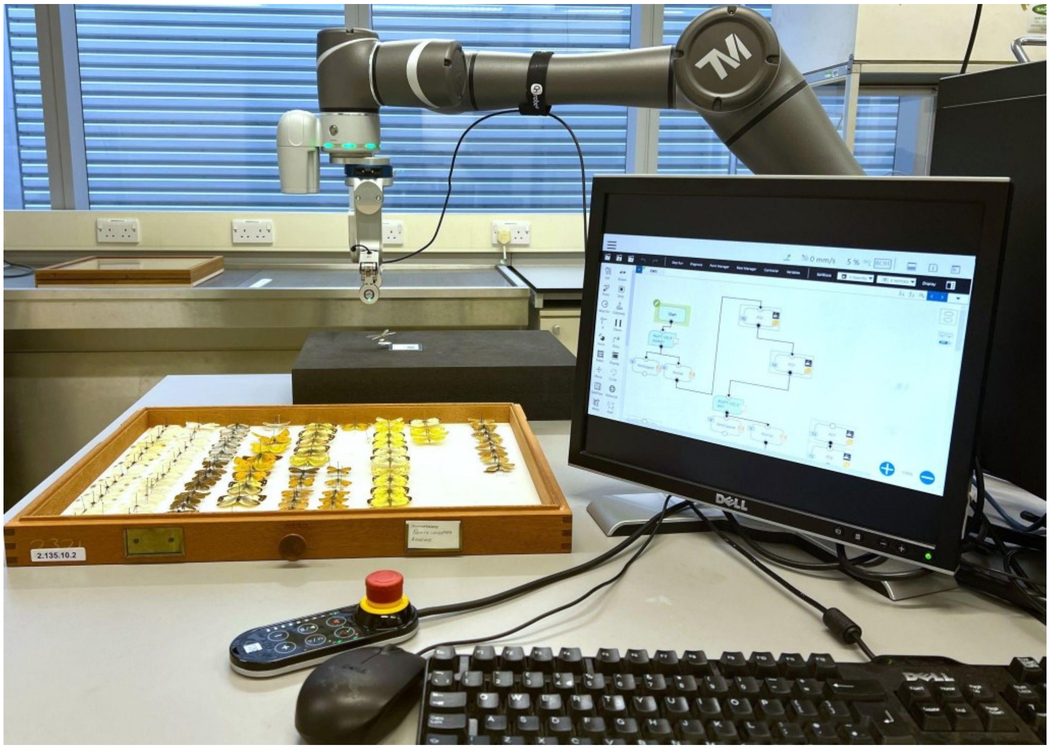

Figure 5: A Techman 500 robotic arm in action at the Natural History Museum, London, placing down a sample pinned specimen from a Lepidoptera collection. Here, the robotic arm has been trained to locate the specimen from the drawer, and then pick it up and place it on a board in order to scan the specimen.

### Example 1b: Label Extraction within Digitisation Pipelines

There is a wealth of data contained within or alongside museum specimens. Specimen labels are a good example of an attribute that is often just as vital in digitisation as the specimen itself. ML tools along with the latest digitisation innovations, have allowed for the development of techniques that enable digitisers to automatically extract information from labels whilst digitising specimens. For example, a cost-effective and efficient pinned insect digitisation process was introduced by Price et al. (2018), which involved placing the specimen within a light box and capturing a handful of photographs simultaneously with multiple cameras from varying angles. The framework described there and in Salili-James et al. (2022b) shows how one can turn to ML to merge labels together from the differently angled images in order to obtain clean, unobstructed images of labels and hence automatically extract textual information from them for digitisation purposes (Figure 6). The first step in this process is reliant on DL tools such as CNNs to locate labels from the multiple images of the specimen. Next, various mathematical and computer vision tools are used to *stitch* the found labels together, in order to have one clear image of each label. These labels can then be fed into an Optical Character Recognition (OCR) and then an NLP algorithm to transcribe the text and to automatically obtain trait information. This leads to a streamlined, automated pipeline to extract label information that helps speed up digitisation efforts.

In general, ML allows for trait extraction to be more easily embedded within digitisation pipelines. One area where this is proving highly effective is in automatic trait extraction from digitised herbarium specimens (Walker et al., 2022), with pilot studies have shown promising results on different types of plants. For example, in LeafMachine (Weaver et al., 2020), a CNN algorithm was trained to measure leaf area and perimeter from low-quality images, with a success rate of 60%. In another study, a different CNN algorithm (ResNet50) was shown to be capable of discriminating between growth shoot and vegetative structures in tropical plants from French Guyana (Goëau et al., 2022). While this study showed a high false positive rate of 20% when identifying growth shoots, it performed well given the complexity and variability of these structures. Overall, these methods have been shown to be able to quickly identify important ecological and evolutionary parameters from herbarium specimens, while still holding a large potential for improvement by expanding the training dataset and refining the algorithms.

As well as the actual plant specimen, herbarium sheets can contain multitudes of data including textual information about the specimen such as location, collector, date, and morphological and colour information if scale and colour bars are included. Text data extraction in particular can be very helpful for digitisation, as the information is embedded with the photograph when digitised onto the database - and this is often done by a digitiser performing manual transcriptions. ML methods can now be used to help speed up the digitisation process of herbarium sheets, for example, by using a combination of different models to automatically extract and categorise textual information during the digitisation process. Another example of ML related to herbarium sheet digitisation involves knowledge graphs. Knowledge graphs (KGs) are an exciting tool in DL that broadly enable a representation of data structured in a graph, with interlinking entities. This allows users to define relationships between different items in datasets. Furthermore, knowledge graphs can then be used to form knowledge bases which can help model large datasets and hopefully one day allow for the creation of a planetary knowledge base (Gu et al., 2023). Beginning with datasets of herbarium sheets, one can build on knowledge graphs (Gu et al., 2022) and create a knowledge base as seen in Gu *et al.* (2023) that can help the digitisation of Herbarium sheets from filling-in missing textual data and to automating transcription tasks.

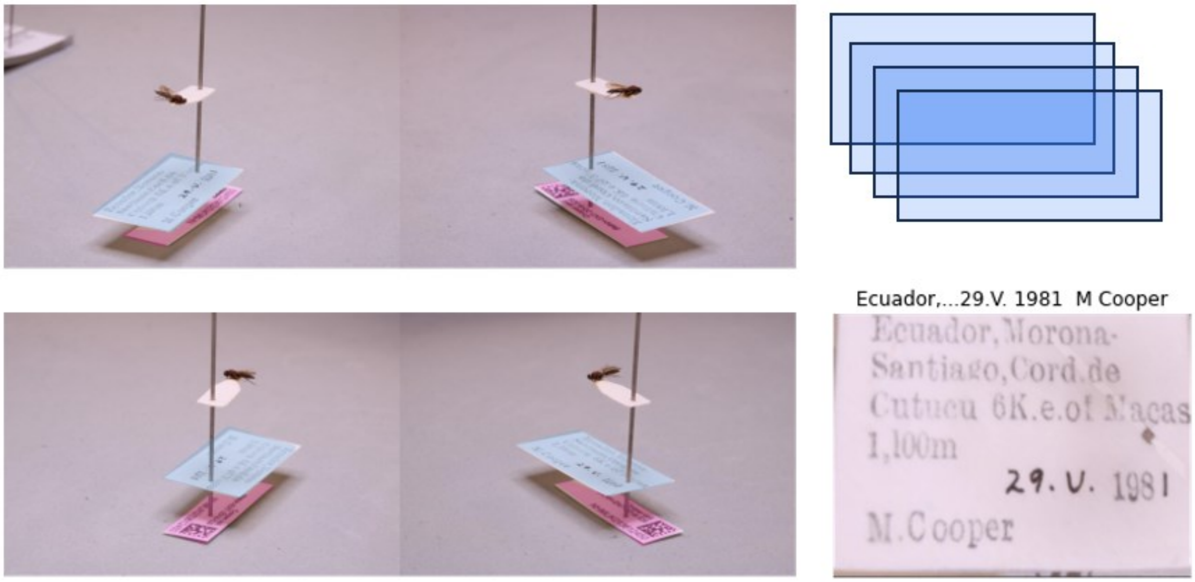

Figure 6: An example of the workflow described in Salili-James *et al.* (2022b). With the ALICE setup introduced in Price *et al.* (2018), the algorithm uses a CNN model to segment all labels found on each of the four images of the specimen. For each label, it then merges the four layers together in order to have one version of each label, which can be fed into an automatic transcription algorithm (with OCR tools). On the bottom right we see an example of a merged label, with a sample of the automatically transcripted text above it.

## Image & scan data collection

While we refer to the use of images for specimen cataloguing above, here we focus on the details of image data collection and analysis. The use of images is central to the study of evolutionary morphology, from simple drawings and photographs to 3D computed tomography (CT) scans (Cunningham et al., 2014). The ability to generate high-resolution images has increased exponentially in recent years, particularly with initiatives for mass-scanning of collections and databases for open sharing of image data (e.g. Phenome10K, Goswami, 2015; Morphosource, Boyer et al., 2016; and DigiMorph, Rowe, 2002). Two-dimensional digitisation of collections often involves photographing collections (i.e., specimens, drawers, etc.), to create digital copies of the data. These photographic images can then undergo segmentation or region identification and extraction, where specific components are identified and separated from the image for further processing or evaluation.

Present-day efforts to digitise specimens with two-dimensional images for large-scale data acquisition and utilisation often involve some automated processes, which can streamline both digitisation and the interpretation of data overall. Recent studies (Salili-James et al., 2022b; Scott and Livermore, 2021) describe software that uses ML models to identify regions of interest in two-dimensional images. Once trained, ML and DL software can capture photographs, segment regions of interest (ROIs), and complete other tasks for large collections datasets. This streamlines the overall acquisition and processing of digital data. Over time, ML software becomes more accurate as it learns through training datasets. This means that efficiency and accuracy can increase as the software is exposed to more data.

The use of automated robotics for digitisation and high-throughput data collection has historically been applied to two-dimensional methods such as photography (e.g., Case Study 1). Three-dimensional data, such as micro-CT data, can also be collected with new robotic technologies like autoloaders (Rau et al., 2021). Autoloaders allow users to set up multiple specimens for micro-CT and synchrotron scanning, set distinct parameters for each scan, and subsequently run the autoloader without supervision. The autoloader processes specimens in a queue, pulling each from the stand using a robotic arm, and setting up distinct parameters for each (Rau et al., 2021). This fully-automated process results in greater efficiency of acquisition, as the number of specimens digitised via this method increases when digitisation can occur without technician supervision. Whilst use of robotic technology to digitise collections could greatly increase the efficiency of image collection, the improvements are more than mechanical. Robots can learn behaviours through Reinforcement Learning (trial and error, as well as rewarding and/or punishing). By interacting with the environment (e.g., the digitisation room), robots can learn optimal actions that maximise rewards (e.g., successfully imaging a specimen).

Novel and potentially more efficient scanning methods are continuously emerging. For instance, Neural Radiance Fields (NeRF) is a fully-connected neural network that can generate a 3D scan of an object by inputting photos of it from different viewpoints (Martin-Brualla et al., 2021). Compared with traditional photogrammetry and CT-scanning, this method is able to compute three-dimensional scans based only on sparse images (Yu et al., 2021). While the resolution and accuracy are typically inferior to a full 3D scan, it can make 3D data capture more accessible and faster for some objects (e.g., extremely large specimens).

## Image Data Processing

Image data capture has become increasingly available in recent years, with large programmes focused on mass scanning of natural history collections (Hedrick et al., 2020). The bottleneck has now shifted to processing images in order to obtain usable data on phenotype. Here, we focus on the two major aspects of image data processing: feature extraction and element isolation, and segmentation.

## Extracting features and isolating elements

Image segmentation refers to dividing an image into meaningful areas or objects and extracting ROIs, allowing for targeted analysis and understanding of visual content (Yu et al., 2023). Segmentation facilitates numerous computer vision tasks, including object recognition by isolating objects or regions within an image (Garcia-Garcia et al., 2018; Jin et al., 2022), object tracking (Zhao et al., 2021), and interpreting a scene with multiple objects (Byeon et al., 2015). This process has traditionally been performed manually; however, it remains subjective (Schmelzle et al., 2017) and time-intensive (Hughes et al., 2022). ML techniques offer a useful way of overcoming these issues.

Numerous automated image segmentation algorithms have been developed in the past decades that do not require DL (Boykov et al., 1999; Dhanachandra et al., 2015; Minaee and Wang, 2019; Najman and Schmitt, 1994; Nock and Nielsen, 2004; Otsu, 1979). Yet, in recent years, DL has introduced novel methods linked to high-performing models able to achieve high accuracy rates on common benchmarks (Kale and Thorat, 2021; LeCun et al., 2015; Luo et al., 2021; Yu et al., 2022; Zhao et al., 2021). DL-based segmentation methods are the state-of-the-art for many image segmentation challenges and often outcompete other automated methods.

For example, Sashimi, a toolkit developed by Schwartz & Alfaro (2021), was introduced to simplify high-throughput organismal image segmentation using DL. The toolkit underwent testing by automatically segmenting the target fish images from photos with both standardised and complex, noisy backgrounds. By utilising advanced DL techniques, such as the meta-algorithm Mask R-CNN, it aims to improve the efficiency and precision of image segmentation, ultimately leading to significant progress in image analysis and classification.

Beyond extracting a complete representation of a specimen from an image, most biological applications will need to identify specific features, whether individual traits or entire elements. Perhaps more than any other aspect, this step is overwhelmingly manual at present and as such represents the primary bottleneck for big data phenomic analyses from comparative datasets. Automated approaches to identifying and isolating features, traits, and elements, however, is perhaps the most important area to develop, as it would potentially allow for continuity of information on biological homology in large-scale comparative analyses and for removal of subjective decisions in trait descriptions. For example, methods that could extract individual elements of a larger structure could then allow one to conduct quantitative analyses of just those isolated structures, or analyses of the relationships among structures (i.e. phenotypic integration and modularity (Zelditch and Goswami, 2021)), where existing automated morphometric methods typically are homology-free and capture overall shape but cannot identify which elements are actually changing in a multi-element structure.

For image data, most efforts at feature extraction have focused on 2D images, extracting features such as size and shape, or using pixel intensity and edge characteristics to extract a plurality of features that are then subjected to dimensionality reduction by identifying those that are most relevant for the particular task. This reduction can select features based on impact of downstream results (e.g., classification), or by using PCA or similar approaches to identify the major components of variation across a dataset (Grys et al., 2017). Along these lines, there are established applications of AI approaches to the recognition of phenotypes and phenomics in medical research, for instance in mining health records (Frey, 2018). These approaches have also been used to identify known phenotypes associated with specific genetic variants (Zhang et al., 2022) or experimental treatments, as in Latent Space Phenotyping, a novel image analysis method that allows the automated recognition of the response to treatments from sequences of images, e.g. different growing conditions of crops (Ubbens et al., 2020). DL has also been applied to feature recognition relating to movement, such as in human foot bone morphology (Ma and Zhi, 2022). These approaches are less common for interspecific datasets; however, they have been applied to recognize and categorise fossil samples. Elsayed (2023) developed an automated approach for identifying and classifying tooth fossils from various animals, including sharks, elephants, hyrax, and primates. Deep-learning algorithms, such as CNN and deep CNN, were trained to recognize, classify, and extract pertinent details from 2D images of teeth fossils. The proposed methods were verified and can be used in various morphology and palaeontology research settings.

## Segmentation

ML and DL approaches to image segmentation are also routinely applied to images forming a 3D stack and achieves accurate results with processing scans from CT (Ait Skourt et al., 2018) and magnetic resonance imaging (MRI) (Lösel and Heuveline, 2017) (Case Study 2). Due to the special characteristics of these types of imaging, such as greyscale and volumetric images with depth information, specific ML and DL models have been developed, particularly for medical images (Milletari et al., 2016; Ronneberger et al., 2015). These methods allow for greater consistency amongst measurements and allow for scalable studies (Willers et al., 2021). In addition, user-friendly tools for segmenting medical images have been developed that offer built-in features for automatic image segmentation such as Dragonfly (Comet Technologies Canada Inc., 2022) and Biomedisa (Lösel et al., 2020). These have since been applied to biological systems (Lösel et al., 2023; Mulqueeney et al., 2024).

---

## Case study 2: Image segmentation for volume rendering

DL tools such as Biomedisa (Lösel et al., 2020) have emerged as powerful solutions for automating feature extraction from 3D images (Figure 7). Additionally, semi-automated batch-processing pipeline MiTiSegmenter can bulk segment and label around 200 samples from microCT data (Kendrick et al., 2022). They offer an efficient alternative to labour-intensive and potentially biased manual image segmentation methods. In the study by Mulqueeney et al. (2024) the efficacy of these neural networks is shown to be influenced by the quality of input data and the size of the selected training set. In the context of this case study, this is reflected in the ability for different networks to extract specific traits. In the smaller training sets, predicting the volumetric and shape measurements for internal structures presents a greater challenge compared to the external structure, primarily due to sediment infill (Zarkogiannis et al., 2020b, 2020a). However, by increasing the training set and applying data augmentation, this problem is mitigated. This reaffirms the principle that expanding the training set leads to the production of better DL models (Bardis et al., 2020; Narayana et al., 2020), albeit with diminishing returns as accuracy approaches 100% (Kavzoglu, 2009). These findings help to highlight how training sets can be designed for optimal use in precise image segmentation that is applicable for obtaining a wide range of traits.

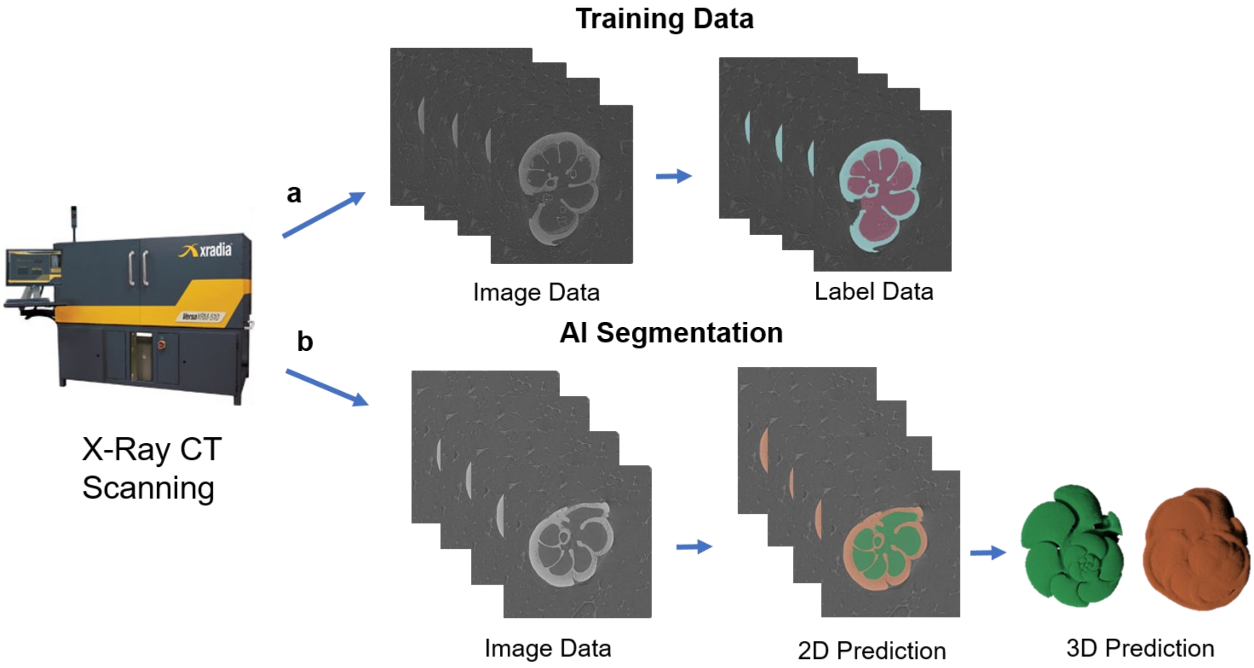

Figure 7: Workflow for producing training data and applying a deep convolutional neural network (CNN) to perform automated image segmentation. The workflow includes (a) the creation of training data for the input into Biomedisa and (b) an example application of the trained CNN to automate the process of generating segmentation (label) data.

---

Beyond increasing efficiency of segmentation over manual thresholding, DL-assisted segmentation may be beneficial whenever thresholding ROIs is not possible. For example, when specimens being scanned are very dense, scans may not have a consistent perceived density (e.g. Alathari, 2015; Furat et al., 2019). Objects of similar densities may not be displayed at the same greyscale value through the scan, though the structural properties of the material will be evident. Scans like these are often also very noisy as a result of the high power of the beam needed to penetrate them, this frequently results in artefacts and irregularities within images (Das et al., 2022). Hence, thresholding cannot always obtain a clear segmentation, leaving manual segmentation as the only recourse prior to these models being implemented. A DL segmentation model however can be trained to segment scans based on visual patterns when a minimal number of slices are pre-labelled (Tuladhar et al., 2020). Noteworthy uses of this approach include distinguishing fossils from rock matrices with a comparable composition within CT images (Edie et al., 2023; Yu et al., 2022), a common problem when imaging palaeontological specimens. In their recent study, Yu et al. (2022) addressed the difficulties of labelling and segmenting CT images, specifically in dinosaur fossils. Traditional manual segmentation methods can be time-consuming and prone to errors due to subjective judgments. To overcome these limitations, the authors employed DL techniques, specifically CNNs, to automate and improve the accuracy of CT image segmentation.

Another case where DL segmentation may be useful for CT data is when attempting to segment regions of an object made of the same material (i.e., if an object of a single material ossifies as a single structure but has varying patterns of ossification along the structure) or when multiple objects have similar densities. DL segmentation models have also shown to accurately segment these objects into different ROIs based purely on pattern. Improvements in the quality of image data acquisition (Withers et al., 2021) alongside the increasing selection of good models (L. Wang et al., 2021) and training sets (Nikolados et al., 2022) for biological data are further assisting in mitigating these common issues in image segmentation with DL approaches.

The ability to easily separate discrete modules has the potential to greatly simplify workflows that are currently mostly manual. An application of these automated or semi-automated segmentation procedures is the isolation of complex biological features or structures that are both time-consuming and difficult to extract manually, such as vertebrate skeletal systems. For instance, skulls are made up of several distinct, overlapping bones–sometimes including other elements such as horns and teeth–the challenge of separating these elements efficiently would be greatly simplified by the application of these methods (Case Study 3). Additionally, this would allow the segmentation of morphological features enclosed within or defined by the interaction between bones, such as endocasts and closed cavities within bones, as well as open-ended sutures. More work in this area will be critical for AI approaches to evolutionary morphology.

Beyond file types, the majority of current methods use human sculpted 3D elements as benchmarks (Chen et al., 2009). These 3D models are generally very low polygon count and manifold, and as a result do not reflect the majority of real-world examples. Work by Schneider *et al*. (2021) attempted to address this by developing a segmentation pipeline able to process higher-polygon and non-manifold meshes. This is ideal for geometric morphometrics, where variations in morphology of focal specimens are only discernible when meshes have sufficient polygons to properly map their topology.

Finally, while identifying known phenotypes from supervised learning is relatively straightforward, although practical limitations of computational power can hinder 3D image data analyses, it is less clear whether unknown or novel phenotypes are similarly recognisable, or if trained models can accommodate large amounts of variation, both of which will be common in analyses of evolutionary morphology. Nonetheless, employing AI to identify new or cryptic features, and by extension, new species, has great potential, particularly in light of promising applications of unsupervised learning to discover unknown phenotypes, for example in cell morphology (Choi et al., 2021).

---

## Case Study 3: Feature extraction and region/specimen isolation via Parcellation

When employing automated shape analysis tools, it is often required to extract features from the entire mesh. For instance, in the context of 3D geometric morphometrics applied to the skulls of antlered animals, it is currently necessary to exclude the antlers and teeth from the specimen, as historical landmarked data has. These regions exhibit significantly higher morphological variation when contrasted with the skulls themselves. Moreover, the antlers and teeth might warrant their own shape analysis, independent from the skull.

Below are outlined the major steps for a case study of segmenting antlers and teeth from the skull (Figure 8). We utilise Blender, an open-source 3D software, along with its Python scripting tools for this purpose. Initially, we import the meshes as PLY files and, optionally, centre them to the origin. If you choose to centre the mesh, it is crucial to export it to ensure that the new coordinates align with the subsequent segmentation. Once centred, access the "scripting" workspace at the top of the window. Run the first script to initialise the vertex groups. Change the 3D view to “edit mode” and the select mode to “edge selection”. Begin selecting the edges of specific classes and assigning them to the associated vertex group. After assigning a cluster of edges run the second script to update the colour of the mesh to match the new vertex groupings. Because vertices can be assigned to more than one group, the antlers and the teeth were assigned first to prevent overlapping specific edges. The third script is then run to assign all unassigned vertices to the skull group, this ensures no edge is assigned to more than one group. Open the system console under the window tab then run the fourth script, this will check that all vertices are assigned and mutually exclusive. Finally, run the script to export the edges.

There are numerous practical issues to solve in extracting features and traits from 3D image data. As in image segmentation, one of the foremost challenges in feature or trait extraction emerges when differentiating tissues or objects exhibiting low contrast disparities, often arising from either similar material densities or the specific imaging techniques employed (Tuladhar et al., 2020). As noted above, conventional methods such as thresholding or region-growing face difficulties in precisely discerning objects under such conditions. To address these limitations, contemporary solutions harness the power of DL. Similarly, DL also finds application in extracting distinct parts of biological anatomy, utilising either semantic segmentation techniques (Hou et al., 2021) or well-crafted training sets (Lösel et al., 2023), even in scenarios where density values closely resemble each other (Case Study 4). Although these methods currently require some manual intervention, they undeniably contribute to substantial reductions in processing times. Challenges may still arise, particularly in the presence of artefacts or irregularities in images (Das et al., 2022) or when dealing with damaged samples (Zhang et al., 2022). Nevertheless, ongoing advancements in these techniques indicate a trend of continual improvement and broader application, and we highlight some recent examples here.

Different AI algorithms have proven capable of classifying taxa from CT images using feature extraction. For example, Hou et al. (2020) introduced the ADMorph dataset, which trained and evaluated DL models for the morphological analysis of 3D digital microfossils. The study focused on enhancing the accuracy of DL models by testing the segmentation performance of multiview convolutional neural networks, PointNet, and VoxNet. The dataset aims to facilitate developing and evaluating DL algorithms for the precise analysis and classification of microfossil structures. Hou et al. (2021) further expand the prior study by highlighting the potential of DL to automate segmentation and accurately delineate and classify about 500 fish microfossils within CT images.

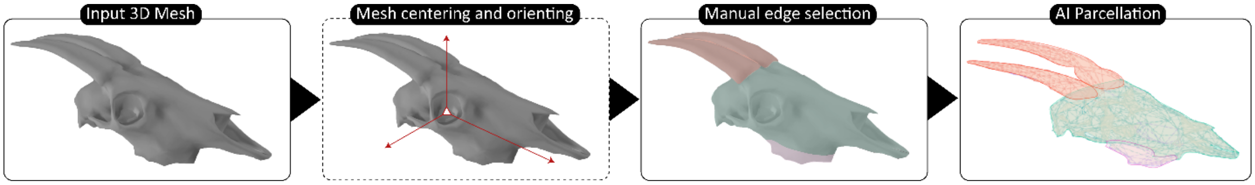

Figure 8: Workflow for segmenting antlers and teeth from a skull using Blender

---

As noted above, parcellation of segmented elements allows for more in-depth analysis of specific areas of focus. In 2D analysis, these methods are present in behavioural ecology and neuroscience, where limb tracking of segmented species in video footage is used to infer behaviour of individuals (Marks et al., 2022; Mathis et al., 2018). Similar to 2D, 3D semantic segmentation using CNNs has started gaining traction, notably in the field of pathology (Rezaeitaleshmahalleh et al., 2023; Schneider et al., 2021), engineering (Bhowmick et al., 2020; Kong and Li, 2018) and materials science (Holm et al., 2020; Zhu et al., 2020), and is similarly useful for evolutionary morphology. For example, extracting individual structures, such as sutures, from micro-CT scans of whole crania allows detailed analysis of their morphology and the factors driving their evolution (Case Study 4).

This approach, however, comes with some important challenges when applied to 3D data. First, the high diversity of data types and extensions in which 3D reconstructions can be stored (e.g., .ply, .vol) poses problems to the homogenisation of automated segmentation procedures. Second, the quality of 3D data can also be an issue. Current methods normally employ human-sculpted 3D elements as benchmarks (Chen et al., 2009), which tend to have low polygon counts and thus do not reflect most biological datasets. As a result, semantic segmentation of 3D reconstructions has proven challenging, with various methods attempting to overcome quality issues in the CT data (Schneider et al., 2021; Shu et al., 2022; Sun et al., 2023).

---

## Case study 4: Image segmentation for feature recognition and automatic trait extraction

Instead of parcellating scans, segmentation can be used directly as phenotypic features. This section focuses on a use case of the segmentation of cranial sutures for phenotypic analysis.

Cranial sutures are fibrous bands of connective tissue that form the joints between the cranial bones of vertebrates (White et al., 2021). These features are bounded on two sides by the bones that they connect, but, unless fused, they are open-ended dorsally and ventrally (Lenton et al., 2005). This presents an issue in that their digital isolation (segmentation) is a highly time-consuming and skill-intensive task, which makes building large datasets for biological comparisons very challenging. We propose a pipeline to address this methodological challenge using DL (Figure 9), summarised as follows:

(1) Create training data: As sutures are difficult and time-consuming to label, we segment one out of every 100 slices (or a specified number of slices) to include as many scans in the training set as possible. Additionally, a test set is created with sutures segmented throughout the entire stack for a few scans, which can ensure the evaluation is more robust.

(2) Train models: We use the training set to train DL models. To address the class imbalance issue, which is caused by sutures normally being small regions, we implemented specific sampling and weighting techniques. We then evaluate the model performance on the test set.

(3) Predict: Sutures for the rest of the scans can be predicted using a well-performing model from the training. These predictions can be reviewed by experts to generate high-quality suture segmentation. The resulting segmentations can be used as a new training set to enhance model performance, or used for downstream analysis.

After this pipeline, we expect to segment sutures from skull CT-scans efficiently. Subsequently, we can use computational methods such as Fourier transform and alpha-shape analysis and landmark-free geometric morphometrics to quantify characteristics of the sutures.

Beyond sutures, such a pipeline would be applicable to segmenting (both in 2D and 3D) any open- or close-ended structure, biological or not, that is defined by the interactions between other structures (i.e., cranial endocasts, chambers in mollusc shells, cracks in bones and other materials, junctions between cells).

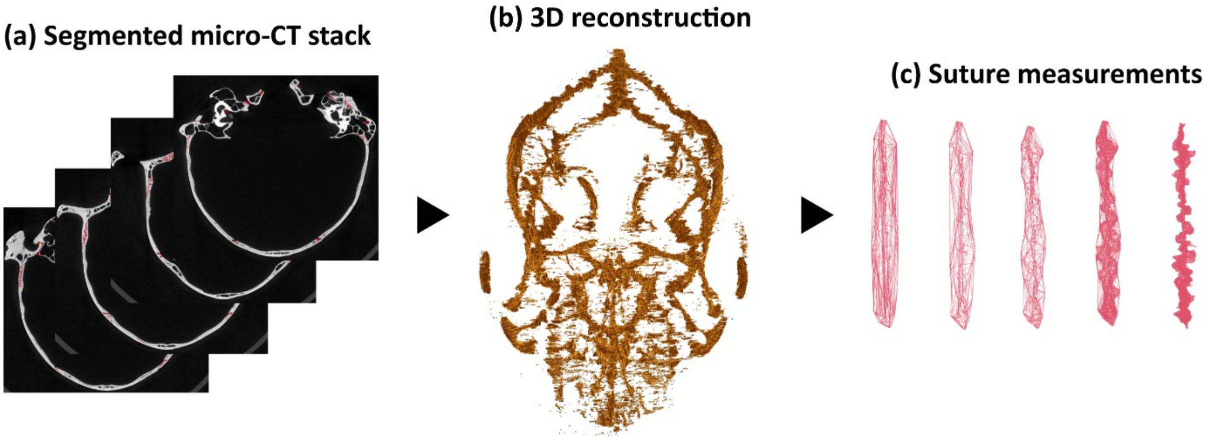

Figure 9: A workflow for extracting sutures on Micro-CT scans. This workflow includes (a) segmenting sutures on Micro-CT scans of mammal skulls. Segmented sutures are used to generate (b) 3D reconstructions, which can then be used to calculate (c) suture measurements. (c) shows an example of using alpha-shapes to quantify the complexity of 3D suture reconstructions following method of Gardiner et *al*. (2018).

# Phenomics

Phenotype encompasses morphology, behaviour, development, and physiology, all of which mediate an organism's interactions with other species and its habitat. Phenomics extends phenotype to its genetic, epigenetic, and environmental drivers. Analysis of phenomes thus entails a variety of traits, all of which are essential to be able to understand the dynamics of organismal evolution, yet the resolution as to which we can currently measure is limited. Here, we discuss how AI techniques can be used to more effectively describe phenotypic traits specific to morphology, with sections related to discrete and meristic traits, univariate measures, shape (including linear and geometric morphometrics), colour, and pose estimation.

## Discrete and Meristic Traits

Morphological traits underpin the study of phenotypic evolution within phylogenetic systematics (Hennig, 1966). Nonetheless, morphological traits for phylogenetic applications have many limitations (Lee and Palci, 2015). Discrete traits manually scored by each researcher or meristic traits such as element counts collected from specimens have proven time-consuming and difficult to collect due to personal interpretations and potential errors (Wiens, 2001). Despite this, discrete traits are critical for diverse aspects of evolutionary study; for example, they are essential to time-calibrate molecular phylogenies and to reconstruct phylogenetic relations among extinct taxa (Lee and Palci, 2015; Smith and Turner, 2005). Discrete and meristic data are also useful for evolutionary analyses of morphology, evidenced by foundational works of morphological disparity (Foote, 1997, 1993; see Goswami and Clavel, 2024 for a full review).

AI tools have shown potential in recognising and extracting discrete and meristic traits to build morphological matrices for phylogenetic analysis in a quicker and more robust way. Deep learning neural networks, including CNNs have been successfully applied on small training datasets to recognise species and extract both discrete and meristic traits (Wäldchen and Mäder, 2018). Other examples include using ML tools to extract, classify and count reproductive structures (Goëau et al., 2022; Love et al., 2021), as well as to produce basic measurements such as leaf size (Hussein et al., 2021; Weaver et al., 2020). These methods have also been shown to work on x-ray scans of fossil leaves (Wilf et al., 2021), including counting stomatal and epidermal cells for palaeoclimatic analysis (Zhang et al., 2023). A similar CNN algorithm has also been successfully applied to classify freshwater fish by genera from the Amazon region using photos of museum specimens, for which, traits were recognised with 97% confidence (Robillard et al., 2023). In animal species traits identification, Random Forest algorithms have also shown promising results. For example, they performed better than traditional Linear Discriminant Analysis in delimiting between species of snakes from field photos when given a set of morphological traits (Smart et al., 2021).

Overall, each of these algorithms have the potential to be used in morphological trait extraction and phylogenetic analysis by training them to classify new images for a set of traits and using a training dataset representing the variation in the species or genus for the traits of interest.

## Univariate Measures

Morphometrics, the quantification of biological form, allows for direct comparison of size and shape across structures or organisms. Univariate metrics have dominated morphometrics for centuries, but the extraction of univariate traits from a substantial pool of individuals has historically been a laborious and time-consuming process, imposing limitations on available data (Fenberg et al., 2016). Addressing this challenge, AI tools have emerged as effective solutions, streamlining the extraction of univariate traits, including lengths, mass, and size, particularly in 2D images. For instance, neural networks have proven adept at extracting linear measurements, as illustrated by the accurate forewing length extraction of 17,000 specimens of butterflies (Wilson et al., 2023). Moreover, these AI techniques have extended their capabilities beyond simple length measures, such as by measuring plant leaf areas (Kishor Kumar et al., 2017; Mohammadi et al., 2021). Advanced techniques have further facilitated the measurement of length across individual anatomical regions, offering a more nuanced understanding than traditional whole-body length measures (Ariede et al., 2023). These techniques have also enabled the extraction of shape proxies, such as ellipticity (Freitas et al., 2023), and the simultaneous analysis of multiple univariate traits (Fernandes et al., 2020).

AI methodologies have seamlessly extended their proficiency from extracting 2D univariate traits to 3D, by employing analogous methods to obtain linear measurements of both length and width within 3D images (Hu et al., 2020; Lu et al., 2023). Similar to the techniques applied to their 2D counterparts, these methods can concurrently extract multiple traits from individual images (Wu et al., 2021) and tally features across diverse regions in 3D images (Yu et al., 2021). Moreover, they are capable of providing volumetric measures of multiple components through the application of image segmentation (Lösel et al., 2023; Mulqueeney et al., 2024). This advancement represents a notable stride in using AI to extract intricate 3D data.

## Shape

Univariate or linear morphometrics has been a tool in evolutionary morphological analysis for centuries, but recent years have seen an explosion of geometric (landmark-based) and surface morphometrics, greatly increasing the scope for capturing and quantifying organismal shape. While surface methods are relatively new, they are expanding rapidly, and geometric methods are well established and offer great potential to increase understanding of evolutionary dynamics (Mitteroecker and Schaefer, 2022). Below we discuss various approaches to multivariate shape analysis and existing AI applications.

### Geometric morphometrics

One of the most common ways to quantify and analyse evolutionary morphology at present is by applying geometric morphometrics. This inherently multivariate methodology requires the placement of landmarks that produce two-dimensional or three-dimensional coordinates by labelling homologous anatomical loci to describe biological shapes (Adams et al., 2004; Mitteroecker and Schaefer, 2022). Raw coordinates are then transformed using a superimposition method, commonly Procrustes analysis, which uses scaling, rotation and transformation to register objects to a common reference frame so that only biological variation remains (Bookstein, 1997). The main advantages of geometric morphometrics include the ability to densely sample complex shapes in three dimensions, the ability to localise variation, the retention of information on biological homology, and the utility of coordinate data for numerous downstream analyses, from macroevolutionary to biomechanical analysis. However, geometric morphometric methods are time-consuming, prone to observer bias, and lack repeatability (Shearer et al., 2017). Moreover, current methods are limited in their scope, in particular because they rely on homologous points of comparison. As a result, they quickly lose explanatory value with increasingly disparate taxa, as homologous points become more difficult to identify and thus fewer in number. As a result, geometric morphometrics is still a largely manual endeavour for comparative datasets, with even semi-automated tools requiring extensive manual effort (Bardua et al., 2019). The introduction of new automated approaches for the analysis of shape may allow us to overcome some of these issues, though the need for grounding in homology will always be a constraint, as well as a critical requirement for maintaining biological meaningfulness, of this approach.

Two distinct groups of approaches have been developed to circumvent the manual effort of geometric morphometrics: automated landmarking and 'landmark-free' or 'homology-free' methods. The former is based on the same principles as geometric morphometrics, but seeks to minimise the user's workload by automating placement of homologous landmarks. These methods frequently rely on image registration to propagate landmarks from one set of individuals, or a generic template, to another (Maga et al., 2017; Young and Maga, 2015). These often lack precision in identifying anatomical loci, even in closely related taxa; therefore, to improve the obtained results, others have attempted to use DL and computer vision to the problem of landmark annotation. In 2D images there is the capacity to automatically place landmarks (Porto and Voje, 2020; Case Study 5), while those available for 3D images at present use AI to optimise landmark position after placement, thereby improving accuracy after mapping of landmarks from a template to specimens (Devine et al., 2020). Landmark placement tools are currently only available for 2D images (Porto and Voje, 2020), but these methods have been shown to reduce both data collection time and error and increase repeatability, thereby supporting phenomic-scale data collection for large data sets. Unfortunately, some implementations are computationally demanding to apply, particularly when using 3D data, and all applications at present behave poorly with even a moderate amount of variation, effectively limiting applications to analysis of conspecifics or congeneric species.

## Case study 5: Geometric morphometrics - automated landmarking

Geometric morphometrics is a powerful tool for quantifying and comparing morphology across organisms and has become the dominant approach for morphometric analysis over the past few decades. However, identifying and placing landmarks and semilandmarks on 2D or 3D images remains a largely manual and thus time-consuming process, as well as being prone to human error and interuser variation. Thus far, AI has been successfully applied to landmark and semilandmark data capture within species and among congeners, ranging from fruit flies (Porto and Voje, 2020; Salifu et al., 2022), to bryozoan colonies (Porto and Voje, 2020), to mice (Devine et al., 2020; Porto et al., 2021). 3D applications at present use image registration and map landmarks from an atlas onto specimens, and then integrate DL for optimization of landmark placement (Devine et al., 2020). Perhaps the most advanced implementation of DL for landmarks placement at present uses a supervised learning approach combining object detection and shape prediction to annotate landmarks (Figure 10) (Porto and Voje, 2020). Object detection, using a histogram of gradients features rather than the more common but less efficient CNN approach, was used to first identify the structure of interest, followed by shape prediction to annotate landmarks. This approach was successfully applied to three datasets of varying complexity, with object detection in particular performing well for all datasets. While only implemented for 2D images at present, the speed of data collection achieved in that study is remarkable (e.g., >13,000 bryozoan zooids annotated in three minutes, approximately the same needed to manually annotate one zooid; Porto and Voje, 2020) and demonstrates the great potential of AI applications to geometric morphometrics and the need to develop implementations for 3D data.

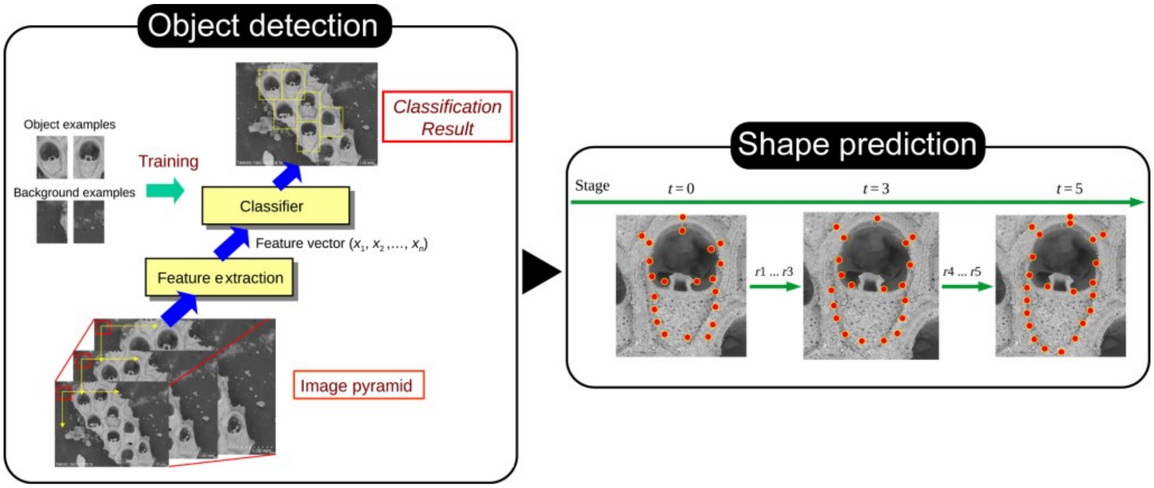

Figure 10: Workflow for automated landmarking in Porto and Voje (2020), showing (a) the object detection framework where a training set is used to first extract features and then perform classification and (b) perform shape prediction using a cascade shape regression model to refine the landmark predictions.

---

## Landmark-free morphometrics

Landmark or homology-free methods seek to remove the placement of landmarks altogether; instead, they focus on describing the entire shape of the selected specimens. There are several methods within this family and most do not directly use AI at present, but we note a few of them as they are promising areas of current development. The most common approaches either decimate a mesh into a large number of pseudolandmarks (i.e., points without any homology) (Boyer et al., 2015; Pomidor et al., 2016) or use an atlas-based diffeomorphic approach (Durrleman et al., 2014; Toussaint et al., 2021). Both of these approaches allow shapes that do not share homology to be compared and limit the loss of geometric information, but they may be prone to sensitivity to factors outside of just shape, including alignment and scaling. Nonetheless, they offer a potentially rich source of data for AI applications, as we discuss here with particular emphasis on diffeomorphic methods.

Broadly, diffeomorphic methods involve a shape on a deformable grid that can be stretched and compressed, with mathematical tools called *diffeomorphisms*, to resemble other shapes. These methods, often referred to as methods of elastic shape analysis due to the *elastic* nature of them, can be used to quantify dissimilarities between shapes, register (*match*) shapes together, and analyse morphometry, all without requiring landmarking. Techniques that incorporate these methods include Large Deformation Diffeomorphic Metric Mapping (LDDMM) (Beg et al., 2005), the Square Root Velocity Framework (Srivastava et al., 2011), and Currents (Benn et al., 2019). One way *elastic* landmark-free techniques are proving increasingly useful is when analysing morphometry in a two-dimensional sense, for example, when studying the boundaries of objects seen in images. Here, instead of requiring landmarks on the boundaries, the boundary curve is analysed as a whole. Importantly, this also allows for possible invariances to be handled. For example, the metrics within methods can be made to be invariant to shape-preserving transformations, such as scaling, translation, rotation, and/or reparametrization (i.e., *where* on the boundary, the curve starts/ends).

There are some diffeomorphic methods that can be expanded into higher dimensions as seen with open curves (Lahiri et al., 2015) and closed curves (Klassen and Srivastava, 2006) - this can prove particularly useful in the analysis of curves on surfaces in evolutionary datasets. There has also been recent research on elastic methods focused on surfaces (Hartman et al., 2023; Jermyn et al., 2017; Pierson et al., 2021). As described in Hartman *et al*. (2023), the techniques here can be categorised into two sections, those that apply to parameterized surfaces and those on unparametrized surfaces (i.e. containing no known point landmarks).

Methods of elastic shape analysis can play an important role in ML, whether this is with classical methods or combined with the latest DL tools. These applications can broadly be split into the following three categories:

1. **Elastic Shape Analysis as a classical machine learning tool**: One motivation for the use of elastic shape analysis is that it allows us to consider the space of the objects we wish to analyse and define our metrics with respect to this space; therefore not assuming linearity of spaces (unlike traditional landmark-based methods). This enables us to create a framework to compute statistics, analogous to standard tools on linear spaces. One such analogue is Principal Component Analysis (PCA), which is a standard tool for dimension reduction in classical ML. This can be done in various ways, from defining ways to work directly on the space of objects or a tangent space to that space (as it is often mathematically and computationally simpler, as described in Srivastava *et al*. (2011)), or by transforming the original objects into a linear space, and then performing standard PCA, as can be done with the framework seen in Benn *et al*. (2019).
2. **Elastic methods as a pre-processing technique for machine learning**: In recent years we have seen elastic diffeomorphic methods applied to real-world datasets and then combined with ML algorithms to provide tools to analyse morphological data. This technique is particularly helpful when a distance matrix can simply be incorporated within a ML algorithm. For example, in Salili-James *et al*. (Salili-James et al., 2022a), diffeomorphic metrics were used to quantify differences between the shapes of natural objects such as boundary curves of gastropods and leaves and then combined with a classical supervised ML algorithm (namely, K-Nearest Neighbour) to classify genus and species, based purely on the morphology of the object.
3. **Machine Learning algorithms with elastic metrics**. ML algorithms will often have a *distance metric* embedded within them. These are especially apparent in classical (statistical) methods of classification (which is SVM) and clustering (which is K-means), where a metric is used to compare the relative position of two points in a dataset. The choice of metric here can sometimes greatly affect the results, and naturally there are

times where a landmark-free elastic (e.g., geodesic) metric can be better suited to the algorithm than standard (often linear) metrics. Furthermore, there have recently been studies that have incorporated DL techniques with elastic metrics, such as in Hartman et al. (2021). Here, a Siamese neural network was trained to predict square root velocity distances between curves, such as the boundary curves of leaves from the notable Swedish Leaf Dataset (Söderkvist, 2016, 2001). Thus, it is justifiable to expect more studies showing DL methods combined with elastic metrics in the near future.

Another family of approaches mathematically describe the surface of the selected objects using integral geometry (L. Wang et al., 2021). These surface descriptors can describe the shape of structures independently and may not suffer from issues of invariance and alignment to the same extent as the landmark-free approaches noted above. However, some applications remain limited, and they have not been widely tested with diverse biological datasets; thus, their efficacy remains to be established. Additionally, all of these approaches have drawn some concerns over ignoring homology (Mitteroecker and Schaefer, 2022), though there is great potential for reintroducing homology by combining these approaches with AI tools for feature or trait extraction, as described above and demonstrated in Case Study 3. These approaches could potentially be used not only to study the shape of specific homologous elements, but also could accelerate studies of modularity and integration (Zelditch and Goswami, 2022), which rely on large sample sizes to assess the relationships among structures, how those relationships reflect genetic, developmental and functional associations among traits, and how they influence the evolution of morphology over shallow to deep time scales.

Thus, despite the attention being paid to new AI techniques and its great potential for automating the quantification of shape, there are at present few applications to datasets above the species-level. The methods remain technical and difficult to apply, due to the need for advanced coding knowledge and access to good hardware such as high-memory GPUs or high-performance computing (HPC) systems. Developments are, therefore, required to make these methods more widely accessible and to allow for greater understanding and addressing of their capabilities and limitations.

## Colour

Colour and patterning are key evolutionary components in taxa as diverse as insects, fishes, birds, and reptiles because of their importance in crypsis, aposematism, mimicry, communication, and sexual selection (Cuthill et al., 2017). Understanding how these patterns evolve is, therefore, crucial for understanding broader evolutionary themes such as natural and sexual selection, convergence, parallel evolution, and character displacement. Colour patterning can help researchers to recognise and discriminate between species and is commonly used in taxonomic, behavioural, and ecological studies. Traditionally, studies have been limited to qualitative descriptions, which has restricted analyses to relatively small sample sizes due to the difficulty of manually comparing large numbers of diverse and complex patterns and colour combinations. Quantitative analyses of colour patterning have become more common in recent years, with important large-scale studies being carried out in birds (Cooney et al., 2019; Dale et al., 2015) and butterflies (Van Der Bijl et al., 2020). Furthermore, automated and semi-automated methods have been developed to segment colour from images (Weller et al., 2022) and to quantify and analyse colour patterns (Maia et al., 2019).

Such methods have helped address the limitations of manually processing data, but there are limits to automated procedures. Automated methods are much faster and less subjective than manual methods for colour segmentation, but are less flexible. Van der Bijl et al. (2020) used a colour profiling approach to assess sexual dimorphism in 369 species of butterflies, using a pixelated image to produce a linear sequence of coordinates containing lightness and colour values. This method is effective but time consuming because each specimen must be photographed, with images manipulated and standardised by hand. As a result, although containing an impressive sample size, the total number of species analysed in this study represents only 2% of the estimated 18,500 extant species of butterflies.

ML offers a potential solution to this bottleneck by combining the capacity to process vast amounts of data and to use large datasets to train and refine its approach. Large image datasets of museum specimens are increasingly being made available for researchers and can act as both training datasets and as comprehensive samples for analysis. ML uses feature extraction and classification to process images in species identification (Wäldchen and Mäder, 2018), and this can be translated into comparing and contrasting colour patterning, by quantifying both spectral (i.e., colour and luminance) and spatial (i.e., the distribution of pattern elements) properties of colour patterns across multiple specimens. With this approach, ML methods can identify individual specimens from photographs, reducing the workload by removing the need to manually process images (Maia et al., 2019). One successful implementation is the analysis of camera trap images, with one study focussing on Serengeti images having a 96% success rate compared with a crowdsourced team of human volunteers (Norouzzadeh et al., 2018). ML has further been used to identify individuals within species of small birds (Ferreira et al., 2020), pandas (Hou et al., 2020), and primates (Guo et al., 2020), based on only minute differences in colour pattern.

Furthermore, the preparation and analysis of data workflows can be greatly improved with the use of AI and some of the most significant progress in this area has been conducted on museum bird specimens. DL methods have been applied to segment and extract plumage from images which greatly enhances the speed at which images can be processed and colour information extracted (He et al., 2022; see Case Study 6). This approach has been taken to even higher levels of specificity by applying pose estimation methods to identify specific points of bird anatomy regions to improve the efficiency of the pipeline for extracting colour information from specimens (He et al., 2023).

---

## Case study 6: Colour

Studying macroevolution of morphological traits often requires extensive measurements from digitised specimens. This can become extremely inefficient in large clades such as in the case of birds where there are an estimated 8,000 extant species globally. Understanding the evolution of avian species, therefore, necessitates extensive datasets of specimens even before factors such as sexual dimorphism have been considered. Several recent studies have employed DL models to assess plumage in birds on a global scale (Cooney et al., 2022; He et al., 2022). DL was applied to segment the whole plumage area from photographs of over 120,000 bird photos across more than 7,500 bird species (Figure 11). The colour and ultraviolet (UV) reflectance of the plumage were extracted from the segmentations. These measurements were then used to study the signalling traits among Passerine birds (> 4,500 species). Findings revealed that UV reflectance is widespread across Passerine birds and is strongly phylogenetically conserved, and the light environment plays a significant role in the evolution of UV reflection (He et al., 2022). Another observation was that both male and female tropical Passerine species tend to be more colourful than their temperate counterparts (Cooney et al., 2022). These findings highlight the potential of AI-generated predictions in biological analyses, addressing questions and hypotheses that could not have been answered previously due to limitations in the efficiency of manual methods.

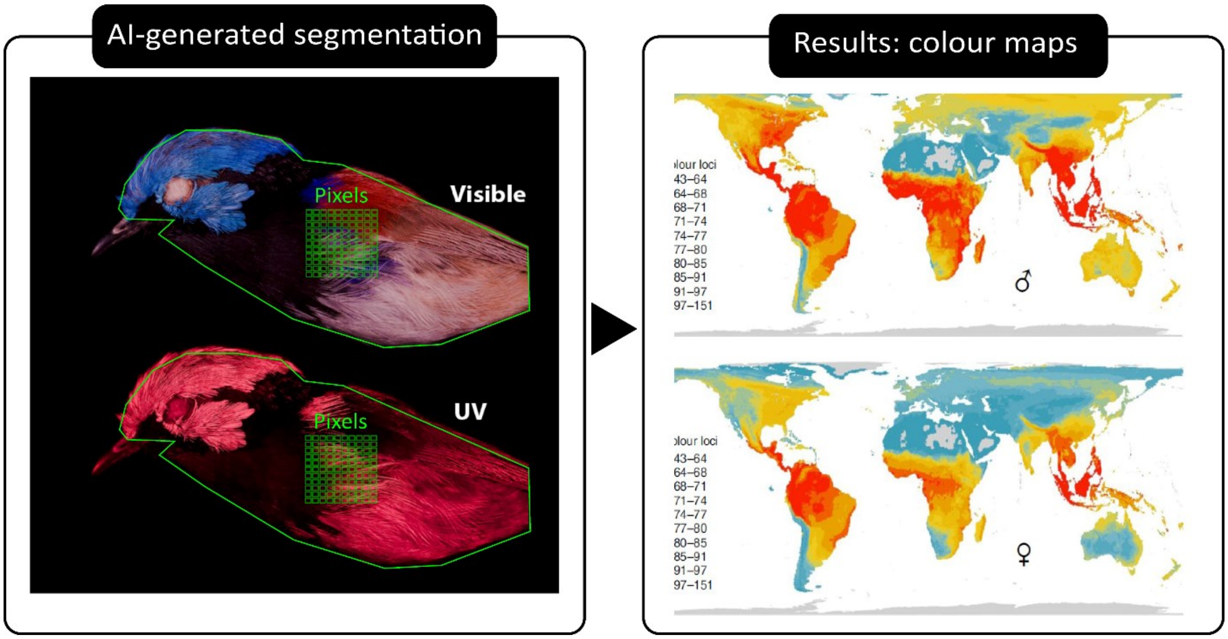

Figure 11: An example of using AI-generated segmentation to study bird plumage colours, modified from He et al. (2022).

## Pose estimation

One of the major avenues of the use of ML, particularly computer vision, in the biological sciences has been pose estimation (Pereira et al., 2019). This approach estimates the relative position of body parts to each other and is used to recognise different animal poses and their changes during locomotion. While estimation is usually conducted on static images, these capabilities have also been adapted to recognise and quantify movement. Indeed, parsing kinematic patterns from videos has become the hallmark of locomotion, biomechanic, and behavioural studies, contributing to the rapid transformation of these fields. Pose estimation is a relatively simple computer vision problem, based on the annotation of training sets from images (Mathis et al., 2018). Originally, algorithms were unable to recognise parts that were not sufficiently distinct from the background, an issue called the 'background problem' (Diaz et al., 2013), and mitigating this required the placement of markers on the moving parts prior to filming. This problem was amplified in video estimation, as motion blur also constituted a significant challenge, requiring the use of extensive and highly specific training datasets (Nath et al., 2019). In light of these issues, the main element of novelty in the field has been the development of computer vision algorithms able to handle video analyses requiring smaller datasets without markers, such as that offered by the recently introduced DeepLabCut toolbox (Mathis et al., 2018; Nath et al., 2019), which has quickly become the standard tool used for marker-free 3D pose estimation (Figure 12). Its capabilities are based on transfer learning: the neural network it is based upon was pre-trained with large datasets, allowing the application of DL to much smaller supervised datasets (Mathis et al., 2018).

Efforts are being made within the field of pose estimation to bridge gaps between biological and computer science expertise. This is increasingly evident in the games and animation industries where there is a need to model animal behaviours for games and films. Manually editing each keyframe can be a painstaking task for animators, thus physics-based models have been employed for years, such as for automatically animating horse gaits (Huang et al., 2013). In recent years, ML tools have been incorporated to automate the process further, such as in the software, WeightShift, which combines full-body physics-based animation with AI to animate characters (Chapman et al., 2020), or in animating the locomotion of quadrupeds using neural networks (Zhang et al., 2018). Another area of pose estimation which has recently benefited from ML is via natural language. AmadeusGPT is a natural language interface for DeepLabCut, which integrates pose estimation and object segmentation (Kirillov et al., 2023). With this the end-user can describe a query and get outputs without needing to code (Ye et al., 2023).

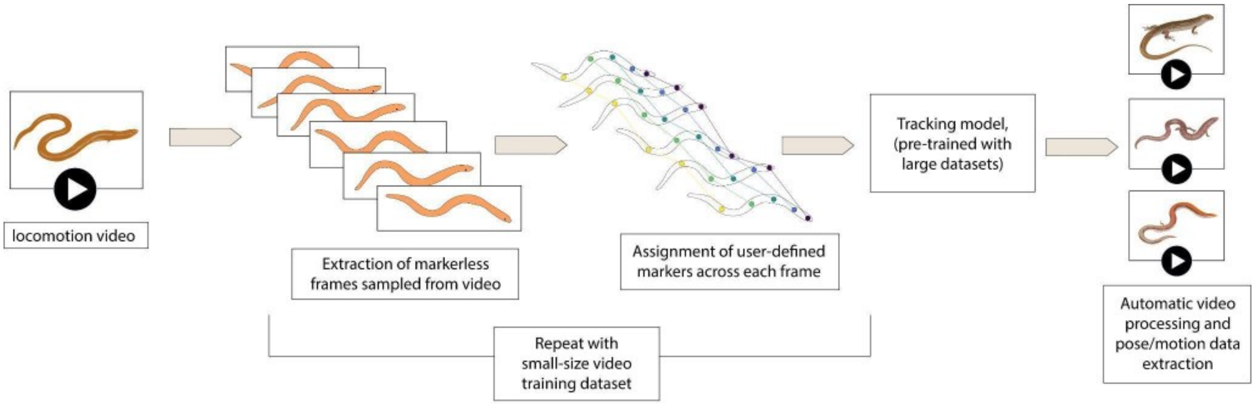

Figure 12: Simplified pipeline for markerless motion tracking and pose estimation from videos using DeepLabCut (Mathis et al., 2018). Limb-reduced skinks (Camaiti et al., 2023) are here used as an example of locomotion tracking.

# Applications to Research in Evolutionary Morphology

AI has the capacity to transform our ability to capture morphology for evolutionary analysis, as detailed above. We are already seeing the implementation of AI approaches for diverse questions in evolutionary biology, but these barely scratch the surface of the potential applications of AI across the field. Below we discuss a range of topics within evolutionary morphology that have already benefited from AI applications, and identify key areas in evolutionary morphology that are ripe for development. In addition to the case studies above that demonstrate how AI is currently being used, we provide a table of tools (Table 1) that are already available for applying AI to evolutionary morphology.

Table 1. Currently available tools using AI for evolutionary morphology

*This table will be regularly updated on [https://phAInomics.github.io/]*

| Tool name / Library | Capabilities | Supported Data types | Programming language | Reference |
|---|---|---|---|---|
| Acquiring Textual Data | | | | |
| NLTK, spaCy<br>*(python libraries)* | Natural language processing (NLP). For example, it can be used for extracting scientific words/taxonomic names from Journal articles | Text | Python | (Bird et al., 2009) |
| TaxoNERD<br>*(python library)* | Extracts scientific names, common names, and name abbreviations.<br>Can link taxa mentioned to a reference taxonomy (e.g. NCBI Taxonomy, GBIF Backbone and TAXREF) | Tabular data, text, images | Python or R | (Le Guillarme and Thuiller, 2022) |

| | | | | |
|---|---|---|---|---|
| pytesseract<br>*(python library)* | Optical character recognition (OCR) to turn images to text. | Images | Python | |
| Google Vision | Deep Learning Application Programming Interface to perform OCR. | Images | N/A | |
| Deep Learning | | | | |
| PyTorch, TensorFlow,<br>*(python libraries)* | DL frameworks. | Tabular data (arrays, matrices etc)<br>Image based data<br>Text<br>Audio | Python | (Martín Abadi et al., 2015; Paszke et al., 2019) |
| Scikit-learn<br>*(python library)* | Tools for classical ML. Classification methods (e.g. Support Vector Machines), clustering methods (e.g., K-means clustering), dimension reduction (e.g. PCA). | A variety of datatypes, from tabular data, to image and sound data etc. | Python | (Pedregosa et al., 2011) |
| PIL, scikit-image, open-cv-python<br>*(python libraries)* | Image processing and computer vision tools. For example, thresholding, contour extraction with Snakes (Active Contour). | Images | Python | (van der Walt et al., 2014) |

| | | | | |
|---|---|---|---|---|
| Monai,<br>Biomedisa<br>*(python libraries)* | DL tools that are designed for processing medical images | Images, especially medical images | Python | (Cardoso et al., 2022; Lösel et al., 2020) |
| Image processing software | | | | |
| ORS Dragonfly,<br>Avizo-Amira,<br>VGSTUDIO MAX | Softwares for processing and segmenting medical and cross-sectional images. AI-based segmentation methods are also supported. | Medical images | The software is not open-source; but it supports Python scripting. | (Dragonfly: Comet Technologies Canada Inc., 2022; Avizo: Thermo Fisher Scientific, 2021) |
| 3D Slicer,<br>Imagej | Open-source softwares for processing medical and cross-sectional images. Users can add extensions such as SlicerMorph, or build their own extensions | Medical images | C++, Python, Qt | (Kikinis et al., 2013; Rolfe et al., 2021; Schneider et al., 2012) |
| Tools can be used in evolutionary morphology | | | | |
| MeshCNN | Mesh classification and segmentation<br>Can be used for segmenting 3D mesh models of specimens | 3D mesh models | Python | (Hanocka et al., 2019) |
| Detectron2 ML library | Object detection | Images | Python | (Wu and Kirillov, 2019) |

| | | | | |
|---|---|---|---|---|
| | Can be used for identifying a specimen in an image. | | | |
| Segment Anything | A pre-trained segmentation tool that can generate decent segmentation results | Images | Python | (Kirillov et al., 2023) |
| Pl@ntNet | Species ID through identification of traits for plants | Images | N/A, input images directly to online tool (identify.plantnet.org) | (“Pl@ntNet IPT,” 2023; “Pl@ntNet,” 2023) |
| FloraIncognita | Species ID and identification of traits for plants | Images | N/A, input images directly to online tool (floraincognita.com) | (Mäder et al., 2021) |
| Fishial.ai | Species ID and feature recognition for fish | Images | N/A input images directly to web portal (portal.fishial.ai) | (“Fishial.ai,” 2019)<br>fishial.ai |

|  |  |  |  |  |
|---|---|---|---|---|
| Merlin Bird ID | Species ID for birds from descriptions, photographs, and sound recordings | Images<br>Audio | N/A, input images directly to mobile app (merlin.allaboutbirds.org) | (Cornell Lab of Ornithology, 2024) |
| Wolfram Mathematica | Identifying type of specimen in an image<br>Categorising traits of specimens from images | Images | Wolfram Language, C/C++, Java | (Wolfram Research, Inc., 2024)<br>Wolfram Research, 2024 |
| MaxEnt | Modelling taxa’s ecological niches | Species occurrence data, environmental rasters | Java | (Phillips et al., 2024) |

## Clustering and classification

Genetic data is commonly used in constructing phylogenies for extant species; however, this is not possible in certain circumstances, including the majority of extinct taxa. Due to these shortcomings, ML techniques using images have become widespread in the classification of individuals into distinct species (Barré et al., 2017; Hsiang et al., 2019; Valan et al., 2019; Wäldchen and Mäder, 2018). Current research predominantly employs CNNs (Krizhevsky et al., 2012), which excel at extracting features from images and providing probability estimates to assign images to specific species classes. These methods, however, only classify the species and do not describe the relationships between classes.

Some AI-based image recognition methods have sought to overcome this issue and possess potential for phylogenetic applications, but studies so far are limited. Kiel (2021) describes a method combining DL and computer vision approaches to train a CNN to categorise images of bivalve species into family groupings based on “known” taxonomy. For each species image, the algorithm estimates the probability that it belongs to one family rather than another. These probability scores are used as a proxy for morphological similarity and to construct a distance matrix, which is in turn used to cluster the families and infer a topology. Whilst this method did find significantly more bivalve families clustering with members of their “known” subclasses than expected by chance, the resulting phylogeny did indicate many “unlikely” placements. When multiple CNNs trained at different taxonomic levels were combined, the resulting phylogeny more closely matched the expected clustering based on existing taxonomic standing.

Morphometric data is also available for use in species identification, and in recent years ML techniques such as Artificial Neural Networks (ANNs) have been employed to accurately classify species (Salifu et al., 2022). These may be a better source of information to reconstruct the evolutionary origins of entire clades as they are able to quantitatively measure the degree of difference amongst individuals. However, like the image classification methods, they rely on supervised training with taxonomic labels to assign species, introducing an inherent assumption of phylogeny.

Each of these techniques must identify distinct morphological attributes for grouping, posing challenges for species-level phylogenies, especially for fossil taxa with limited individual samples. Lastly, using morphological similarity as a proxy for phylogenetic placement, without a specified model, may be prone to the effects of homoplasy and convergent evolution. Despite these constraints, the ability to use ML algorithms to differentiate taxa based on morphology may be useful to infer relatedness based on morphological similarity/dissimilarity.

## Species delimitation

Species delimitation, opposed to classification, requires the ability to identify whether individuals belong to a population, which in some cases may lead to new species being introduced. The limitations of image clustering methods makes them ill-suited for this task. Instead emerging techniques in one-class classification systems (Perera and Patel, 2019) or open set recognition (Geng et al., 2021) offer promising avenues for extending species identification beyond initial classifications done through image analysis. However, inherent challenges remain; these techniques are currently used for outlier detection and would need to be adapted to establish species.

An alternative approach would be to use phenotypic traits as a basis for delimitation. Individuals can be grouped into self-similar clusters by analysing phenotypic traits, forming the basis for delineating populations and species (Ezard et al., 2010). Traditionally, Gaussian mixture models (GMMs) employing a Maximum Likelihood approach have been utilised (Fraley and Raftery, 2002). However, the advent of deep Gaussian mixture models (Viroli and McLachlan, 2019), which incorporate ML techniques, may be more suitable. These models show heightened levels of complexity, enabling them to capture intricate relationships within data. These approaches, combined with the increasing ability to acquire image or trait data rapidly, may allow for a more nuanced and comprehensive understanding of taxonomy.

A number of genomic species delimitation methods have been extensively used in the last decade, including Bayesian species delimitation (BPP), which has been cited over 600 times (Yang, 2015). In addition, unsupervised ML algorithms have been employed  on genomic data to predict clusters of individuals (Derkarabetian et al., 2019). More recently, convolutional neural networks have been utilised to build a morphology-molecule network [MMNet] that integrates morphological and molecular data for species identification (Yang et al., 2022). However, despite their widespread adoption and increasing applications in taxonomy, these methods do not work when faced with species that are not present in the training set, which renders them ineffective for identifying novel or undiscovered species.

Nonetheless, DL tools can be used to detect anomalies both in labels and in species classification. From a practical level, this can be used to assist curators in adding correct information (as with the Knowledge Bases; Gu et al., 2023), and even one day detect new species. One application is to use a well-trained classifier model to classify taxonomic information (e.g., genus or species) of specimens that are being digitised. Then by comparing the predictions and actual specimen labels, potential missing or incorrect labels can be flagged. Similarly, this approach has been applied to the classification of species from digitised specimens, as in Hansen *et al.* (2020) where CNNs were used to classify beetles from images of digitised museum specimens. These models, particularly when combined with further classification and clustering tools, such as with heatmap analysis (Hollister et al., 2023), can one day be used to identify new species by simply scanning digitised museum collections.

Similar to the accumulation of image data, many more genomic markers and whole genomes are being sequenced today. Consequently, unsupervised or semi-supervised AI-based integrative taxonomic tools have the potential to play a key role in furthering species discovery. In addition to phenotypic traits and genomes, researchers are obtaining additional suites of organismal data such as acoustics, behaviour, and ecology. AI will be key to bringing these complex datasets together for a biologically meaningful interpretation of a ‘species’.

## Phylogenies - Building Trees

Evolutionary studies frequently involve tree data structures, especially phylogenetic trees. The use of machine learning (ML) methods remains relatively uncommon in phylogenetic inference, and research into this area is still in its infancy. Despite recent progress and potential to address limitations of traditional methods, the extent of ML's accuracy and scalability remains uncertain. ML algorithms can mitigate the computational costs associated with Maximum Likelihood and Bayesian approaches, are flexible, and do not necessarily require explicit specification of models. However, a significant obstacle for supervised methods is the scarcity of training data for tree inference. A ‘true’ phylogeny is fundamentally unknowable, leading to reliance on simulated data that may not accurately reflect evolutionary relationships. Recent reviews of ML approaches for tree building (Mo et al., 2023; Sapoval et al., 2022) have predominantly focused on molecular phylogenetics, leaving morphology-based phylogenetics relatively unexplored despite its unique opportunities (e.g. incorporating fossils) and unresolved challenges. Currently available models of morphological evolution are generally more simplistic than those that exist for molecular evolution (Lee and Palci, 2015), though are likely more complex in reality. However, ML could be applied to developing morphological models of evolution through automated assessment of trait covariations (which could also be applied to studies of modularity and integration, as noted above), changes through time using existing phylogenies, and probabilities of key innovations versus gradual variations. This is a key area for development; at present there are no published attempts to apply AI methods to morphology-based phylogenetic inference, or to build models to estimate morphological change through time to aid in building those phylogenies.

## Molecular phylogenetics

There are currently far more studies focused on applying ML methods to genetic data than to morphological data. However, a number of approaches being tested using sequence data may have potential for future application to morphological data. CNNs and RNNs have been employed to infer quartet (4 taxa) topologies using simulated sequence alignments and protein data (Suvorov et al., 2020; Zou et al., 2020). These methods can be trained to deal with instances of extreme model violation where traditional statistical methods such as Maximum Likelihood might fall short, and once trained, tree estimation can be very fast (Zaharias et al., 2022). Simulated quartet experiments have been shown to outperform methods like Maximum Likelihood, with particular success in scenarios of high substitution heterogeneities, which many standard models struggle to account for (Zou et al., 2020). However, more recent analyses contest this, and traditional methods have outperformed neural network methods when the taxon number is increased above four (Zaharias et al., 2022).

These methods have mostly been applied to individual sequences, but applying them to species trees involves further complexities such as incomplete lineage sorting and introgression (Degnan and Rosenberg, 2009; Maddison and Knowles, 2006; Suvorov et al., 2020). Restrictions of limited taxa and the complexity of species tree inference are emerging areas of research, such as in a recent study applying generative adversarial networks (GANs) to simulated data and seven species of fungi (Smith and Hahn, 2023). The proposed phyloGAN model uses two networks: a generator that suggests new topologies, and a discriminator trained to differentiate real and generated data, effectively deciding how “realistic” a proposed topology and alignment might be. This method imitates the heuristic search employed by many traditional methods to explore tree space for more optimal trees. PhyloGAN shows an improvement in the number of taxa that can be considered compared to previously mentioned methods, but is still limited compared to traditional methods, and hampered by lengthy computational times (Smith and Hahn, 2023).

Another molecular ML tree building approach is Phyloformer which computes distances between molecular sequences in a multiple sequence alignment (MSA) (Nesterenko et al., 2022). This method simulates trees, then uses probabilistic models of sequence evolution, working backwards to simulate MSAs. Supervised learning is then used to train a ML algorithm to reverse engineer the phylogeny based on an associated MSA. In the case of Phyloformer, the algorithm estimates pairs of evolutionary distances between sequences that can then be used to infer a tree using traditional methods such as Neighbour Joining (NJ). Phyloformer was found to outperform standard distance-based methods, and also performed competitively against Maximum Likelihood whilst being significantly faster.

The methods described here use different degrees of ML to estimate evolutionary relationships, either by approximating distances between taxa or by directly inferring topologies. It is not yet clear whether they will be applicable to morphological data. Methods such as Phyloformer still rely on models of sequence evolution. Such models are lacking in the field of morphometrics due to increased complexity and the lack of clearly defined smallest units of change across the tree of life. This presents a challenge to the application of this framework of methods, although morphological data could be coded into such analyses as matrices.

### Model selection, heuristic searches, and missing data

One of the most important considerations for phylogenetics is the type of model used. However, identifying the optimal model can be challenging. ML algorithms have recently been applied to improve model selection methods. ModelTeller (Abadi et al., 2020) and ModelRevelator (Burgstaller-Muehlbacher et al., 2023) are two such approaches that focus specifically on identifying the most appropriate substitution models for a particular analysis or dataset. Whilst both focus on molecular substitution models, their existence opens the possibility of developing new systems for selecting morphological evolutionary models.

Many phylogenetic methods (including Maximum Likelihood and Bayesian) employ heuristic searches, where model parameters (such as tree topology and branch length) are adjusted and the likelihood calculated for each adjustment. This method essentially explores tree space for a set number of iterations, aiming to identify parameter combinations of increasingly higher likelihoods. Such methods are limited by how extensive their tree search is, and as a result can become extremely computationally expensive. ML methods are now being applied to improve the efficiency of this process by predicting which neighbouring trees will increase the likelihood without actually calculating the value, thereby reducing computational expense (Azouri et al., 2023, 2021).

Finally, a major challenge in both molecular and morphological phylogenetic studies is the impact of missing data. This is particularly impactful for distance-based methods where calculating a distance matrix is complicated by the presence of missing data in the alignments. In the case of molecular phylogenetic studies, this refers to missing bases in sequences. For morphological data this could be a result of incomplete specimens where certain traits or biological structures are missing or difficult to measure or score. Previous studies have shown that missing data negatively affects the accuracy of tree inference methods (Roure et al., 2013; Wiens, 2006). ML methods such as PhyloMissForest (Pinheiro et al., 2022), which uses a Random Forest approach, and two methods proposed by Bhattacharjee & Bayzid (2020), use ML to estimate missing distance values within a distance matrix and may outperform traditional statistical methods.

## Phylogenetic comparative methods and evolutionary modelling

Using a phylogenetic framework to estimate the evolution of clades and traits has become a core part of evolutionary morphology over the past few decades (Adams and Collyer 2019; Felsenstein 1985). Analysis of trait variation across phylogenies and through time relies on the availability of well-supported topologies and time calibration. Recent advances in genome sequencing and big-data approaches to taxonomic sampling and trait data collection have increased the availability of time-calibrated phylogenies. In turn, this has enhanced our ability to reliably map the evolution of traits on phylogenies and consider phylogenetic relations when examining relationships between traits across multiple taxa.

The potential applications of AI in reconstructing trait evolution using a phylogenetic framework have been documented theoretically. These applications include character evolution model choice and the reliable and efficient encoding of phylogenetic information. For example, Ho et al. (2019) demonstrated the theoretical application of ML to ancestral estimation of phenotypic traits using a multi-task learning approach applied to Brownian Motion models of continuous biological traits. A multi-task learning approach is an ML framework that pools data from different tasks (Ruder, 2017). This takes the form of an algorithm able to reduce the variance of estimators, using a penalty term that draws models closer to each other, allowing the estimation of ancestral states of multiple characters simultaneously. This approach improved ancestral estimations compared with Maximum Likelihood models at the expense of a slight bias introduced in the phylogenetic estimates (Ho et al., 2019).

Despite theoretical advances, there are currently no practical applications of ML approaches to estimate trait evolution. A known issue that would benefit from an AI-based modelling approach is the assignment of distinct rates of character evolution to different parts of a given phylogenetic tree (i.e., King and Lee, 2015). ML would enable the simultaneous pooling of multiple data sources, including distributions of states at the tips of phylogenetic trees, branch lengths, node ages, uncertainty in node resolution, and hidden states, and consideration of a wide variety of complex models that may better reflect phenomic datasets (Goswami and Clavel, 2024). ML approaches could also facilitate the comparison of simulations across trees. Furthermore, AI methods could account for phylogenetic relatedness in analyses of trait correlations. In the field of bioinformatics, using DNN and Convolutional Graph Network (CGN) architectures in phylogenetic profiling for protein interactions improved predictions (Moi and Dessimoz, 2022). In particular, combining CGN with a graphical representation of tree topology allowed for prediction across multiple species and could be used to predict pairwise interaction across time. Using these deep neural network algorithms in conjunction with phylogenetic information is currently exploratory but could potentially streamline and improve multiple aspects of estimating trait evolution and ancestral states, allowing better modelling of the complex factors underlying evolution on a phenomic scale.

## Function and Adaptive landscapes

In evolutionary biology, adaptive landscapes are conceptual frameworks that illustrate the relationship between the phenotype of an organism and its fitness within a specific ecological context (Arnold, 2003; McGhee, 1999, 1980; Simpson, 1984). They provide a visual representation of natural selection-driven trait space across the blanket of an adaptive landscape, where peaks of specific traits reflect higher fitness compared to putative trait space across the landscape. Over evolutionary time, genetic variation, mutation, recombination, and natural selection drive the population towards regions of higher fitness. Utilising models of trait diversification can be helpful in tracing adaptive peaks of species through time, adapting to different ecological niches or responding to environmental shifts. The study of adaptive landscapes is key both to understanding the evolutionary adaptive mechanisms giving rise to biodiversity and predicting the future adaptive potential of species in light of anthropogenic-driven habitat loss and climate change.

Functional adaptive landscape analysis (FEA) uses the morphology and function of skeletal elements to model landscapes (Dickson and Pierce, 2019; Jones et al., 2021; Polly et al., 2016; Tseng et al., 2023). In palaeontology, functional adaptive landscapes commonly employ FEA as a functional metric (Deakin et al., 2022; Polly et al., 2016). ML algorithms can replace FEA to predict the behaviour of a beam in a one-dimensional system if the algorithms are first trained on initial FEA. ANNs have been suggested to provide more accurate FEA results than boosting regression trees or Random Forest ML algorithms (Vurtur Badarinath et al., 2021). Furthermore, ML algorithms could assist in the morphometrics (Baylac et al., 2003; Punyasena et al., 2012) used in adaptive landscapes in addition to modelling adaptive landscapes through evolutionary time.

Additionally, AI has been increasingly applied to FEA-based biomechanical modelling (Galbusera et al., 2020; Mouloodi et al., 2021). These techniques can be applied to data extracted from static images, 3D-image data (Galbusera et al., 2020), and even motion capture (Mouloodi et al., 2021). The isolation of distinct features (image segmentation) and the capture of locomotory information (e.g. through pose estimation) can both be automated through computer vision, which allows the identification of parts of interest and their spatial relationships to one another. Once features have been extracted, AI can be used to simulate their behaviour in relation to one another under a given set of physical constraints (e.g. Liu, 2019). This is particularly useful for the creation of models of the range of appendicular motion, relationships between internal organs, and even models of cytokinesis (Huiskes and Hollister, 1993; Ross, 2005; Shi et al., 2010).

## Phenome-environment and ecometrics

One of the most established areas of phenotypic analysis is quantification of relationships between phenomes of organisms (the sum of their phenotypic traits) and the environmental context in which they evolved. The end goal of many studies using this approach is to assign an ecomorphological characterisation to phenotypic traits and to parse their ecological signal (Barr, 2018). AI has been implemented in this field through the use of algorithms that infer present and past ecomorphologies by reducing the dimensionality of ecomorphological data through ML pipelines such as Random Forest analyses (Mahendiran et al., 2022; Rabinovich, 2021; Sosiak and Barden, 2021; Spradley et al., 2019). Similarly, ML procedures have been used to discriminate and sort phenotypes (especially morphology) based on their belonging to specific ecomorphs or ecological guilds (MacLeod et al., 2022). These studies have highlighted the advantages of AI-based approaches compared to standard procedures used to test the links between morphology and ecology, such as Canonical Variate Analysis (Albrecht, 1980).

The related field of ecometrics is a taxon-free approach to quantifying the distribution of functional traits across space and time (Eronen et al., 2010). Ecometric correspondence between environmental and phenotypic data is used to develop transfer functions which can be used to reconstruct paleoenvironments or incorporate Species Distribution Modelling (SDM) to model future spatial distributions of phenotypes given predicted climatic scenarios (Parker et al., 2023; Vermillion et al., 2018). Existing work uses linear and maximum likelihood approaches to ecometric modelling. These approaches have a limit of one or two climate inputs, normally limiting analyses to consider only annual precipitation and mean annual temperature (Parker et al., 2023). However, a Random Forest approach would enable the model to use any number of climatic variables. Similarly, SDMs can be built using CNNs, capturing non-linear transformations across multiple variables (Botella et al., 2018). DL approaches to quantifying phenome-environment would enable models to better approach the complex factors contributing to climate and trait distribution, as in studies of trait evolution.

## Niches and Niche Evolution

ML algorithms, including Boosted Regression Tree and Random Forest, have become standard methodologies for modelling the ecological niches of taxa and, by extension, their potential spatial distribution. Over the past decade, research has extensively focussed on predicting the ecological effects of climate change by using ecological niche modelling (Deb et al., 2020; Karuppaiah et al., 2023; Qin et al., 2017; Tang et al., 2021). The most prominent ML model in this area is the ‘maximum entropy modelling method (MaxEnt), which has been applied in thousands of studies since its description in 2006 (Phillips et al., 2006; Merow et al., 2013).

MaxEnt’s ubiquity in scientific literature is in part due to the algorithm requiring relatively few inputs (only species occurrences and geographic data) and relying on biologically reasonable assumptions. It assumes that a taxon will occupy as large an area as possible (maximum distribution entropy; Elith et al., 2011; Phillips et al., 2006). These limitations have also produced an abundance of literature critiquing and subsequently optimising MaxEnt’s statistical assumptions and processes (Campos et al., 2023; Cobos et al., 2019; Low et al., 2021; Sillero and Barbosa, 2021).

Studies that use MaxEnt or other ML methods tend to consider niches as static entities, with many publications ‘projecting’ the same niche onto environmental rasters representing distinct points in time, sometimes thousands or millions of years ago (Saupe et al., 2019). Niche evolution studies have instead relied on measuring the contemporary niche overlap of different taxa (usually via the methodology of Broennimann et al., 2012), considering the similarities and differences within a phylogenetic context (Doré et al., 2023; Padilla-García et al., 2023; Vasconcelos et al., 2023). While both approaches are useful in understanding ecological evolution across time, they are limited by their discrete temporal sampling – niches change continuously across space and time, and an individual niche of a taxon may also change over time.

ML methods could be developed to identify and accommodate niches changing over time. Taxon occurrences sometimes have associated temporal metadata, which could be used by an AI tool to predict the continuous changes in a niche in the recent past or near future. This could prove especially invaluable in studying the effects of climate change at a higher resolution. Considering a geological timescale, the predicted ecological niches of fossil taxa (modelled with environmental data representing periods in deep time) could be used to calibrate and, thus, further validate continuous niche evolution models across phylogenetic trees.

# Prospectus

The scope of evolutionary biology is immense, involving the history of life on Earth over the past >3 billion years. For the vast majority of species that ever lived, the only available data is morphological in nature; thus, studying morphology is crucial for understanding the evolution of organisms. Yet, methods for capturing morphological data remain largely manual, presenting a bottleneck for the study of morphological evolution, particularly in comparison to other biological fields with mature methods for 'Omics' level analyses. The use of AI is bringing about a massive transformation in the field of evolutionary morphology, both for data capture and analysis. Integrating AI techniques into this area will become increasingly important as the field continues to move towards larger-scale analyses and bigger data.

As we have discussed, AI has been successfully applied to a range of data acquisition for evolutionary morphology, and AI applications are only increasing in the pace of development and accessibility for non-experts. For example, AI is already making it quicker to generate, refine, and access image data of larger quantities and/or greater resolutions than ever before. Large gaps remain, however, including discriminating features or regions of interest, extracting discrete traits or 3D morphometric data in datasets with large amounts of variation (which are common in comparative evolutionary analysis), and in applying AI for improving evolutionary models for morphological data. These areas should be the focus of efforts over the coming years. While we have detailed applications of AI to several research areas involving morphological evolution, there are many more for which AI has yet to make a significant impact. Below, we note a few subfields of evolutionary morphology that have clear pathways for improvement through AI. Finally, we close with some considerations on the accessibility and environmental effects of AI.

# Emerging fields

***Retrodeformation* -** Several studies have demonstrated that fossil data are critical for accurately estimating phenotypic evolution through deep time (Slater et al. 2012; Goswami and Clavel, 2024 and references therein). A common challenge in palaeontology is encountering fossils which have undergone taphonomic distortion via brittle or plastic deformation (Kammerer et al., 2020; Schlager et al., 2018). This can severely hamper attempts to assess and quantify intra- and interspecific shape by introducing non-biological variation, and for this reason, as well as the lack of integration in phylogenetic analyses as noted above, fossil data are often excluded from comparative analyses. Retrodeformation is the process of restoring the original shape of an object by reversing this taphonomic distortion (Herbst et al., 2022; Lautenschlager, 2016). While landmark- and symmetry-based procedures to manually perform these operations are available (e.g. Morpho, Schlager et al., 2018), they are time-consuming and can only be applied to relatively small datasets, restricting the taxonomic breadth of studies. AI provides an opportunity to automate and enhance this process. ML models, such as neural networks, can be trained to recognize and correct specific types of deformations. These models can learn patterns of distortion and apply appropriate corrections. In the future, AI may aid in the reconstruction of 3D objects or scans of distorted or even completely flattened fossils, helping to recover valuable 3D morphology. Once models have been trained on a dataset of naturally distorted fossils and manually performed retrodeformation simulations, they can be integrated into software applications or embedded in hardware systems for real-time correction and analysis. The choice of AI techniques and algorithms will depend on the specific application and the nature of the deformations to be corrected. For instance, de Oliveira Coelho (2015) used logistic model trees to predict the temperature at which human bone was burnt. Likewise, Zeng *et al.* (2021) used a support vector machine algorithm to detect small geological faults. Such methods could be co-opted to estimate the extent of brittle and ductile deformation a fossil has undergone, which then enables evolutionary morphologists to apply the opposite forces to correct the distortion.

***Histology*** - Histology examines the microscopic structure and morphology of tissues, including fossil tissues in palaeohistology. Palaeohistology has historically informed on growth, physiology, and development, but is also increasingly used to investigate the form and function of tissues (e.g. the function of duck-billed dinosaur dental batteries; Bailleul et al., 2019, 2012). AI tools have greatly advanced histology, especially in the sub-field of histo-pathology, to better cancer recognition and clinical oncology (reviewed in Shmatko et al., 2022). AI has been proposed to be a promising avenue to increase throughput for pattern recognition. Some areas where AI pattern recognition has already been explored in biological research are in quality assessment of histological images (Haghighat et al., 2022) and herbivore diet characterisation through micro-histological analysis (Filella et al., 2023). Additionally, DNNs have been used to identify primary and secondary osteon regions and create segmented maps of different osteon regions. This osteon segmentation was combined with phylogenetics to elucidate the developmental pathway towards miniaturisation in the theropod dinosaurs Alvarezsauria (Z. Qin et al., 2022). There is significant potential for the use of AI in histological studies in the context of evolutionary morphology and is therefore an avenue for future exploration.

***Genome-phenome mapping*** - AI has been applied in two main areas of genome phenome association (GPA): the medical sciences, and food production. This is not surprising, as both are umbrella areas of research with high societal impact. Deep matrix factorization (DMF)-based methods developed to handle multi-omics data have been successfully applied to improve genome-wide mapping and genome-wide association studies in the context of molecular phenomes. These include disease susceptibility (Long et al., 2023; Mieth et al., 2021), drug efficiency (Mongia and Majumdar, 2020), crop phenotype prediction (Islam et al., 2023), and microbe-disease association (Y. Liu et al., 2021). More recently, a multi-omics data fusion-based approach (Weighted Deep Matrix Genome Phenome Association) has been proposed to incorporate different kinds of omics data and to predict potential nonlinear GPAs with functional traits (Tan et al., 2022). However, for non-model organisms and evolutionary research, AI is underutilised either through GPAs or genome-wide association studies (GWAS). These fields are rapidly evolving with the increasing amounts of data collected worldwide, and therefore provide promising avenues for implementing new AI algorithms in the near future.

***Evo-devo*** - ML has been successfully applied to the study of gene expression in embryonic development of model organisms (Čapek et al., 2023; Feltes et al., 2018; Naert et al., 2021). Algorithms have also been developed to aid in phenotyping and staging embryos and to recognize diseases and malformations (e.g. Al-Saaidah et al., 2017; Jeanray et al., 2015). In evolutionary developmental biology (evo-devo), phenotype identification is a recent development. A few pilot studies have been conducted using both images and morphometrics data on human cells, model organisms and plants (Cai and Ge, 2017; Chen et al., 2020;

Masaeli et al., 2016). CNNs have been used to extract visual patterns from images, to aid embryo staging, and to analyse changes in phenotype during ontogeny (Feltes et al., 2018; Naert et al., 2021).

## Accessibility and Considerations

Until very recently, most AI models were built and applied using Python libraries such as Caffe, TensorFlow, and PyTorch (Jia et al., 2014; Martín Abadi et al., 2015; Paszke et al., 2019), requiring both AI and programming knowledge. Additionally, running these models required specialised, expensive hardware, such as GPUs, which are commonly used in training AI models. Consequently, the required level of expert understanding of AI and costly hardware restricted the accessibility of AI for many researchers in the biological sciences.

As AI continues to advance, it is becoming increasingly accessible to non-experts and more affordable to implement due to several factors. (i) Increasingly user-friendly software has reduced the need for in-depth AI-related knowledge. (ii) The growth of open-source and pre-trained models has significantly reduced the computational resources, data, and time required to develop AI models. (iii) The advent of cloud-based AI services has allowed researchers to access powerful AI without investing in local GPUs. In addition, the cost of robotic arms decreases annually (Zhang et al., 2022), meaning that large datasets capturing phenotype with high resolution are increasingly available.

Despite these advancements, there are certain aspects that require a degree of caution. AI outputs are derived from the data used for training. If the data is biased or unrefined, it could lead to similarly inaccurate and biased results (Mehrabi et al., 2021; Zhang et al., 2022), therefore, attention must be given to data cleaning and preprocessing.

Additionally, the environmental impact of AI cannot be overlooked, particularly as many studies in our fields aim to protect the natural world and limit human-caused climate change and destruction of biodiversity. Evolutionary morphology studies increasingly involve the collection and storage of large quantities of image data. These datasets are currently limited by the hours of manual input required, but will only increase in size as AI approaches allow for more efficient processing and analysis, leading to larger, more complex studies that in turn require increased hardware and energy input. Training large-scale models can consume substantial amounts of energy, contributing to carbon emissions, although admittedly the models trained and used in evolutionary biology are unlikely to be as large as those from tech giants like Google, Meta and OpenAI. Some studies using large scale genetic datasets have estimated the carbon footprint of their computational analyses (Philippe et al., 2019; Y. Qin et al., 2022). More formal approaches to sustainable computer science are being developed in the form of emission calculation tools (Lacoste et al., 2019; Lannelongue et al., 2021), assessments of their suitability for various approaches (Bouza et al., 2023), and proposed principles for greener computational science in the future (Lannelongue et al., 2023). As the scale of AI models and the demand for AI continue to grow, it will be increasingly important for us to evaluate the environmental impact of future studies in evolutionary morphology.

To conclude, we have here provided an introduction and overview of the current and potential future applications of AI to evolutionary morphology. As AI becomes more accessible and tailored towards applications central to the study of evolutionary biology, we expect that it will transform the study of evolutionary morphology. By accelerating and improving capture and analysis of “Big Data” on phenotype for diverse comparative datasets, AI will allow the realisation of evolutionary phenomics and launch a new phase in the study of past and present biodiversity.

# Acknowledgements

For thought-provoking and valuable conversations that have broadened our thinking, we thank Katie Collins, Tom Ezard, and the members of the AI and Innovation group at the Natural History Museum.

# Funding statement

This work was supported by Leverhulme Trust grant  RPG-2021-424 to AG, MC, EG, and YH; Natural Environmental Research Council grants NE/S007210/1 and NE/P019269/1 to JMM; NE/S007229/1 to ECW, NSB, and JM; NE/S007415/1 to ESEH and OKC; BBSRC grant BB/X014819/1 to AG and LER; Lateinamerika-Zentrum Zürich (Switzerland) to GR-deL; EU Horizon 2020 Marie Skłodowska-Curie Actions to AVM; a Daphne Jackson Research Fellowship funded by the Anatomical Society to VH; UKRI grant EP/Y010256/1 to AK and TW, and funding from NHS-X, GSK, and Ely-Lilly to EG.

# Data availability statement

No new data were generated or analysed in support of this research. The tools table in this paper will be kept updated at [phenomeAI.org].

# Competing interests statement

We declare no competing interests.

# References

Abadi, S., Avram, O., Rosset, S., Pupko, T., Mayrose, I., 2020. ModelTeller: Model Selection for Optimal Phylogenetic Reconstruction Using Machine Learning. Molecular Biology and Evolution 37, 3338–3352. https://doi.org/10.1093/molbev/msaa154

Adams, D.C., Rohlf, F.J., Slice, D.E., 2004. Geometric morphometrics: Ten years of progress following the ‘revolution.’ Italian Journal of Zoology 71, 5–16. https://doi.org/10.1080/11250000409356545

Ahmed, K., Keskar, N.S., Socher, R., 2017. Weighted Transformer Network for Machine Translation. https://doi.org/10.48550/arXiv.1711.02132

Ait Skourt, B., El Hassani, A., Majda, A., 2018. Lung CT Image Segmentation Using Deep Neural Networks. Procedia Computer Science 127, 109–113. https://doi.org/10.1016/j.procs.2018.01.104

Akçakaya, M., Yaman, B., Chung, H., Ye, J.C., 2022. Unsupervised Deep Learning Methods for Biological Image Reconstruction and Enhancement: An overview from a signal processing perspective. IEEE Signal Processing Magazine 39, 28–44. https://doi.org/10.1109/MSP.2021.3119273

Alathari, T., 2015. Feature extraction in volumetric images. University of Southampton.

Albrecht, G.H., 1980. Multivariate Analysis and the Study of Form, with Special Reference to Canonical Variate Analysis. Am Zool 20, 679–693. https://doi.org/10.1093/icb/20.4.679

Al-Saaidah, B., Al-Nuaimy, W., Al-Taee, M., Young, I., Al-Jubouri, Q., 2017. Identification of tail curvature malformation in zebrafish embryos. ICIT 2017 - 8th International Conference on Information Technology, Proceedings 588–593. https://doi.org/10.1109/ICITECH.2017.8080063

Amalfitano, D., Faralli, S., Hauck, J.C.R., Matalonga, S., Distante, D., 2024. Artificial Intelligence Applied to Software Testing: A Tertiary Study. ACM Comput. Surv. 56, 1–38. https://doi.org/10.1145/3616372

Angermueller, C., Pärnamaa, T., Parts, L., Stegle, O., 2016. Deep learning for computational biology. Molecular Systems Biology 12, 878. https://doi.org/10.15252/msb.20156651

Ariede, R.B., Lemos, C.G., Batista, F.M., Oliveira, R.R., Agudelo, J.F.G., Borges, C.H.S., Iope, R.L., Almeida, F.L.O., Brega, J.R.F., Hashimoto, D.T., 2023. Computer vision system using deep learning to predict rib and loin yield in the fish *Colossoma macropomum*. Animal Genetics 54, 375–388. https://doi.org/10.1111/age.13302

Arnold, S.J., 2003. Performance Surfaces and Adaptive Landscapes. Integrative and Comparative Biology 43, 367–375. https://doi.org/10.1093/icb/43.3.367

Atz, K., Grisoni, F., Schneider, G., 2021. Geometric deep learning on molecular representations. Nat Mach Intell 3, 1023–1032. https://doi.org/10.1038/s42256-021-00418-8

Audagnotto, M., Czechtizky, W., De Maria, L., Käck, H., Papoian, G., Tornberg, L., Tyrchan, C., Ulander, J., 2022. Machine learning/molecular dynamic protein structure prediction approach to investigate the protein conformational ensemble. Sci Rep 12, 10018. https://doi.org/10.1038/s41598-022-13714-z

Azouri, D., Abadi, S., Mansour, Y., Mayrose, I., Pupko, T., 2021. Harnessing machine learning to guide phylogenetic-tree search algorithms. Nat Commun 12, 1983. https://doi.org/10.1038/s41467-021-22073-8

Azouri, D., Granit, O., Alburquerque, M., Mansour, Y., Pupko, T., Mayrose, I., 2023. The tree reconstruction game: phylogenetic reconstruction using reinforcement learning. https://doi.org/10.48550/ARXIV.2303.06695

Baevski, A., Auli, M., 2019. Adaptive Input Representations for Neural Language Modeling.

Bailleul, A.M., Hall, B.K., Horner, J.R., 2012. First Evidence of Dinosaurian Secondary Cartilage in the Post-Hatching Skull of Hypacrosaurus stebingeri (Dinosauria, Ornithischia). PLoS ONE 7, e36112. https://doi.org/10.1371/journal.pone.0036112

Bailleul, A.M., O'Connor, J., Schweitzer, M.H., 2019. Dinosaur paleohistology: review, trends and new avenues of investigation. PeerJ 7, e7764. https://doi.org/10.7717/peerj.7764

Bardis, M., Houshyar, R., Chantaduly, C., Ushinsky, A., Glavis-Bloom, J., Shaver, M., Chow, D., Uchio, E., Chang, P., 2020. Deep Learning with Limited Data: Organ Segmentation Performance by U-Net. Electronics 9, 1199. https://doi.org/10.3390/electronics9081199

Bardua, C., Felice, R.N., Watanabe, A., Fabre, A.-C., Goswami, A., 2019. A practical guide to sliding and surface semilandmarks in morphometric analyses. Integrative Organismal Biology 1, obz016. https://doi.org/10.1093/iob/obz016

Bardua, C., Wilkinson, M., Gower, D.J., Sherratt, E., Goswami, A., 2019. Morphological evolution and modularity of the caecilian skull. BMC Evol Biol 19, 30. https://doi.org/10.1186/s12862-018-1342-7

Barr, A.W., 2018. Ecomorphology, in: Croft, D.A., Su, D.F., Simpson, S.W. (Eds.), Methods in Paleoecology, Vertebrate Paleobiology and Paleoanthropology. Springer International Publishing, Cham, pp. 339–349. https://doi.org/10.1007/978-3-319-94265-0_15

Barré, P., Stöver, B.C., Müller, K.F., Steinhage, V., 2017. LeafNet: A computer vision system for automatic plant species identification. Ecological Informatics 40, 50–56. https://doi.org/10.1016/j.ecoinf.2017.05.005

Baylac, M., Villemant, C., Simbolotti, G., 2003. Combining geometric morphometrics with pattern recognition for the investigation of species complexes: GEOMETRIC MORPHOMETRICS, PATTERN RECOGNITION and SPECIES COMPLEXES. Biological Journal of the Linnean Society 80, 89–98. https://doi.org/10.1046/j.1095-8312.2003.00221.x

Beg, M.F., Miller, M.I., Trouvé, A., Younes, L., 2005. Computing Large Deformation Metric Mappings via Geodesic Flows of Diffeomorphisms. Int J Comput Vision 61, 139–157. https://doi.org/10.1023/B:VISI.0000043755.93987.aa

Benn, J., Marsland, S., McLachlan, R.I., Modin, K., Verdier, O., 2019. Currents and finite elements as tools for shape space. J Math Imaging Vis 61, 1197–1220. https://doi.org/10.1007/s10851-019-00896-x

Bhattacharjee, A., Bayzid, Md.S., 2020. Machine learning based imputation techniques for estimating phylogenetic trees from incomplete distance matrices. BMC Genomics 21, 497. https://doi.org/10.1186/s12864-020-06892-5

Bhowmick, S., Nagarajaiah, S., Veeraraghavan, A., 2020. Vision and Deep Learning-Based Algorithms to Detect and Quantify Cracks on Concrete Surfaces from UAV Videos. Sensors 20, 6299. https://doi.org/10.3390/s20216299

Bird, S., Klein, E., Loper, E., 2009. Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit.

Blagoderov, V., Kitching, I., Livermore, L., Simonsen, T., Smith, V., 2012. No specimen left behind: industrial scale digitization of natural history collections. ZK 209, 133–146. https://doi.org/10.3897/zookeys.209.3178

Bookstein, F.L., 1997. Landmark methods for forms without landmarks: morphometrics of group differences in outline shape. Medical Image Analysis 1, 225–243. https://doi.org/10.1016/S1361-8415(97)85012-8

Borowiec, M.L., Dikow, R.B., Frandsen, P.B., McKeeken, A., Valentini, G., White, A.E., 2022. Deep learning as a tool for ecology and evolution. Methods Ecol Evol 13, 1640–1660. https://doi.org/10.1111/2041-210X.13901

Botella, C., Joly, A., Bonnet, P., Monestiez, P., Munoz, F., 2018. A Deep Learning Approach to Species Distribution Modelling, in: Joly, A., Vrochidis, S., Karatzas, K., Karppinen, A., Bonnet, P. (Eds.), Multimedia Tools and Applications for Environmental & Biodiversity

Informatics. Springer International Publishing, Cham, pp. 169–199. https://doi.org/10.1007/978-3-319-76445-0_10

Bouza, L., Bugeau, A., Lannelongue, L., 2023. How to estimate carbon footprint when training deep learning models? A guide and review. Environ. Res. Commun. 5, 115014. https://doi.org/10.1088/2515-7620/acf81b

Boyer, D.M., Puente, J., Gladman, J.T., Glynn, C., Mukherjee, S., Yapuncich, G.S., Daubechies, I., 2015. A New Fully Automated Approach for Aligning and Comparing Shapes. The Anatomical Record 298, 249–276. https://doi.org/10.1002/ar.23084

Boyer, D.M., Gunnell, G.F., Kaufman, S., McGeary, T.M., 2016. Morphosource: archiving and sharing 3-D digital specimen data. Paleontol. Soc. Pap. 22, 157–181. https://doi.org/10.1017/scs.2017.13

Boykov, Y., Veksler, O., Zabih, R., 1999. Fast approximate energy minimization via graph cuts, in: Proceedings of the Seventh IEEE International Conference on Computer Vision. Presented at the Proceedings of the Seventh IEEE International Conference on Computer Vision, pp. 377–384 vol.1. https://doi.org/10.1109/ICCV.1999.791245

Breiman, L., 2001. Random Forests. Machine Learning 45, 5–32. https://doi.org/10.1023/A:1010933404324

Broennimann, O., Fitzpatrick, M.C., Pearman, P.B., Petitpierre, B., Pellissier, L., Yoccoz, N.G., Thuiller, W., Fortin, M., Randin, C., Zimmermann, N.E., Graham, C.H., Guisan, A., 2012. Measuring ecological niche overlap from occurrence and spatial environmental data. Global Ecology and Biogeography 21, 481–497. https://doi.org/10.1111/j.1466-8238.2011.00698.x

Brown, T.B., Mann, B., Ryder, N., Subbiah, M., Kaplan, J., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D.M., Wu, J., Winter, C., Hesse, C., Chen, M., Sigler, E., Litwin, M., Gray, S., Chess, B., Clark, J., Berner, C., McCandlish, S., Radford, A., Sutskever, I., Amodei, D., 2020. Language Models are Few-Shot Learners.

Buda, M., Maki, A., Mazurowski, M.A., 2018. A systematic study of the class imbalance problem in convolutional neural networks. Neural Networks 106, 249–259. https://doi.org/10.1016/j.neunet.2018.07.011

Burgstaller-Muehlbacher, S., Crotty, S.M., Schmidt, H.A., Reden, F., Drucks, T., Von Haeseler, A., 2023. ModelRevelator: Fast phylogenetic model estimation via deep learning. Molecular Phylogenetics and Evolution 188, 107905. https://doi.org/10.1016/j.ympev.2023.107905

Byeon, W., Breuel, T.M., Raue, F., Liwicki, M., 2015. Scene labeling with LSTM recurrent neural networks, in: 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Presented at the 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Boston, MA, USA, pp. 3547–3555. https://doi.org/10.1109/CVPR.2015.7298977

Cai, Z., Ge, S., 2017. Machine learning algorithms improve the power of phytolith analysis: A case study of the tribe Oryzeae (Poaceae). Journal of Systematics and Evolution 55, 377–384. https://doi.org/10.1111/jse.12258

Camaiti, M., Evans, A.R., Hipsley, C.A., Hutchinson, M.N., Meiri, S., de Oliveira Anderson, R., Slavenko, A., Chapple, D.G., 2023. Macroecological and biogeographical patterns of limb reduction in the world's skinks. Journal of Biogeography 50, 428–440. https://doi.org/10.1111/jbi.14547

Campos, J.C., Garcia, N., Alírio, J., Arenas-Castro, S., Teodoro, A.C., Sillero, N., 2023. Ecological Niche Models using MaxEnt in Google Earth Engine: Evaluation, guidelines and recommendations. Ecological Informatics 76, 102147. https://doi.org/10.1016/j.ecoinf.2023.102147

Candès, E., Deamante, L., Donoho, D., Ying, L., 2006. Fast Discrete Curvelet Transforms. Multiscale Modeling & Simulation 5, 861–899. https://doi.org/10.1137/05064182X

Čapek, D., Safroshkin, M., Morales-Navarrete, H., Toulany, N., Arutyunov, G., Kurzbach, A., Bihler, J., Hagauer, J., Kick, S., Jones, F., Jordan, B., Müller, P., 2023. EmbryoNet: using deep learning to link embryonic phenotypes to signaling pathways. Nature Methods 20. https://doi.org/10.1038/s41592-023-01873-4

Carbonell, J.G., Michalski, R.S., Mitchell, T.M., 1983. Machine Learning: A Historical and Methodological Analysis. AI Magazine 4, 69–69. https://doi.org/10.1609/aimag.v4i3.406

Cardoso, M.J., Li, W., Brown, R., Ma, N., Kerfoot, E., Wang, Y., Murrey, B., Myronenko, A., Zhao, C., Yang, D., others, 2022. Monai: An open-source framework for deep learning in healthcare. arXiv preprint arXiv:2211.02701.

Chan, T., Vese, L., 1999. An Active Contour Model without Edges, in: Nielsen, M., Johansen, P., Olsen, O.F., Weickert, J. (Eds.), Scale-Space Theories in Computer Vision, Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp. 141–151. https://doi.org/10.1007/3-540-48236-9_13

Chapman, D., Daoust, T., Ormos, A., Lewis, J., 2020. WeightShift: Accelerating Animation at Framestore with Physics. Eurographics/ ACM SIGGRAPH Symposium on Computer Animation - Showcases 2 pages. https://doi.org/10.2312/SCA.20201219

Chen, L., Pan, X.Y., Guo, W., Gan, Z., Zhang, Y.H., Niu, Z., Huang, T., Cai, Y.D., 2020. Investigating the gene expression profiles of cells in seven embryonic stages with machine learning algorithms. Genomics 112, 2524–2534. https://doi.org/10.1016/j.ygeno.2020.02.004

Chen, L.-C., Papandreou, G., Schroff, F., Adam, H., 2017. Rethinking Atrous Convolution for Semantic Image Segmentation. https://doi.org/10.48550/arXiv.1706.05587

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd Acm Sigkdd International Conference on Knowledge Discovery and Data Mining. pp. 785–794.

Chen, X., Golovinskiy, A., Funkhouser, T., 2009. A benchmark for 3D mesh segmentation. ACM Trans. Graph. 28, 1–12. https://doi.org/10.1145/1531326.1531379

Choi, H.J., Wang, C., Pan, X., Jang, J., Cao, M., Brazzo, J.A., Bae, Y., Lee, K., 2021. Emerging machine learning approaches to phenotyping cellular motility and morphodynamics. Phys. Biol. 18, 041001. https://doi.org/10.1088/1478-3975/abffbe

Cobos, M.E., Peterson, A.T., Barve, N., Osorio-Olvera, L., 2019. kuenm: an R package for detailed development of ecological niche models using Maxent. PeerJ 7, e6281. https://doi.org/10.7717/peerj.6281

Comet Technologies Canada Inc., 2022. Dragonfly.

Cooney, C.R., He, Y., Varley, Z.K., Nouri, L.O., Moody, C.J.A., Jardine, M.D., Liker, A., Székely, T., Thomas, G.H., 2022. Latitudinal gradients in avian colourfulness. Nature Ecology & Evolution 6, 622–629. https://doi.org/10.1038/s41559-022-01714-1

Cooney, C.R., Varley, Z.K., Nouri, L.O., Moody, C.J.A., Jardine, M.D., Thomas, G.H., 2019. Sexual selection predicts the rate and direction of colour divergence in a large avian radiation. Nat Commun 10, 1773. https://doi.org/10.1038/s41467-019-09859-7

Cornell Lab of Ornithology, 2024. Merlin Bird ID.

Cortes, C., Vapnik, V., 1995. Support-vector networks. Mach Learn 20, 273–297. https://doi.org/10.1007/BF00994018

Cunningham, J.A., Rahman, I.A., Lautenschlager, S., Rayfield, E.J., Donoghue, P.C.J., 2014. A virtual world of paleontology. Trends in Ecology & Evolution 29, 347–357. https://doi.org/10.1016/j.tree.2014.04.004

Cuthill, I.C., Allen, W.L., Arbuckle, K., Caspers, B., Chaplin, G., Hauber, M.E., Hill, G.E., Jablonski, N.G., Jiggins, C.D., Kelber, A., Mappes, J., Marshall, J., Merrill, R., Osorio, D., Prum, R., Roberts, N.W., Roulin, A., Rowland, H.M., Sherratt, T.N., Skelhorn, J., Speed, M.P., Stevens, M., Stoddard, M.C., Stuart-Fox, D., Talas, L., Tibbetts, E., Caro, T., 2017. The biology of color. Science 357, eaan0221. https://doi.org/10.1126/science.aan0221

Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection, in: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05). Presented at the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), pp. 886–893 vol. 1. https://doi.org/10.1109/CVPR.2005.177

Dale, J., Dey, C.J., Delhey, K., Kempenaers, B., Valcu, M., 2015. The effects of life history and sexual selection on male and female plumage colouration. Nature 527, 367–370. https://doi.org/10.1038/nature15509

Das, S., Nayak, G.K., Saba, L., Kalra, M., Suri, J.S., Saxena, S., 2022. An artificial intelligence framework and its bias for brain tumor segmentation: A narrative review. Computers in Biology and Medicine 143, 105273. https://doi.org/10.1016/j.compbiomed.2022.105273

de Oliveira Coelho, J.P.V., 2015. Unwarping heated bones: a quantitative analysis of heat-induced skeletal deformations using 3D geometric morphometrics (Dissertation). University of Coimbra, Coimbra, Portugal.

Deakin, W.J., Anderson, P.S., den Boer, W., Smith, T.J., Hill, J.J., Rücklin, M., Donoghue, P.C., Rayfield, E.J., 2022. Increasing morphological disparity and decreasing optimality for jaw speed and strength during the radiation of jawed vertebrates. Science Advances 8, eabl3644.

Deb, J.C., Forbes, G., MacLean, D.A., 2020. Modelling the spatial distribution of selected North American woodland mammals under future climate scenarios. Mammal Review 50, 440–452. https://doi.org/10.1111/mam.12210

Degnan, J.H., Rosenberg, N.A., 2009. Gene tree discordance, phylogenetic inference and the multispecies coalescent. Trends in Ecology & Evolution 24, 332–340. https://doi.org/10.1016/j.tree.2009.01.009

Derkarabetian, S., Castillo, S., Koo, P.K., Ovchinnikov, S., Hedin, M., 2019. A demonstration of unsupervised machine learning in species delimitation. Molecular Phylogenetics and Evolution 139, 106562. https://doi.org/10.1016/j.ympev.2019.106562

Devine, J., Aponte, J.D., Katz, D.C., Liu, W., Vercio, L.D.L., Forkert, N.D., Marcucio, R., Percival, C.J., Hallgrímsson, B., 2020. A Registration and Deep Learning Approach to Automated Landmark Detection for Geometric Morphometrics. Evol Biol 47, 246–259. https://doi.org/10.1007/s11692-020-09508-8

Devlin, J., Chang, M.-W., Lee, K., Toutanova, K., 2019. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding.

Dhanachandra, N., Manglem, K., Chanu, Y.J., 2015. Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm. Procedia Computer Science 54, 764–771. https://doi.org/10.1016/j.procs.2015.06.090

Diaz, R., Hallman, S., Fowlkes, C.C., 2013. Detecting Dynamic Objects with Multi-view Background Subtraction, in: 2013 IEEE International Conference on Computer Vision. Presented at the 2013 IEEE International Conference on Computer Vision (ICCV), IEEE, Sydney, Australia, pp. 273–280. https://doi.org/10.1109/ICCV.2013.41

Dickson, B.V., Pierce, S.E., 2019. Functional performance of turtle humerus shape across an ecological adaptive landscape. Evolution 73, 1265–1277. https://doi.org/10.1111/evo.13747

Doré, M., Willmott, K., Lavergne, S., Chazot, N., Freitas, A.V.L., Fontaine, C., Elias, M., 2023. Mutualistic interactions shape global spatial congruence and climatic niche evolution in Neotropical mimetic butterflies. Ecology Letters 26, 843–857. https://doi.org/10.1111/ele.14198

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., Houlsby, N., 2021. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale.

Dumiak, M., 2008. Book-Scanning Robots Digitize Delicate Texts. IEEE Spectr. 45, 18–18. https://doi.org/10.1109/MSPEC.2008.4428300

Durrleman, S., Prastawa, M., Charon, N., Korenberg, J.R., Joshi, S., Gerig, G., Trouvé, A., 2014. Morphometry of anatomical shape complexes with dense deformations and sparse parameters. NeuroImage 101, 35–49. https://doi.org/10.1016/j.neuroimage.2014.06.043

Dutia, K., Stack, J., 2021. Heritage connector: A machine learning framework for building linked open data from museum collections. Applied AI Letters 2. https://doi.org/10.1002/ail2.23

Edie, S.M., Collins, K.S., Jablonski, D., 2023. High-throughput micro-CT scanning and deep learning segmentation workflow for analyses of shelly invertebrates and their fossils: Examples from marine Bivalvia. Front. Ecol. Evol. 11, 1127756. https://doi.org/10.3389/fevo.2023.1127756

Elith, J., Phillips, S.J., Hastie, T., Dudík, M., Chee, Y.E., Yates, C.J., 2011. A statistical explanation of MaxEnt for ecologists: Statistical explanation of MaxEnt. Diversity and Distributions 17, 43–57. https://doi.org/10.1111/j.1472-4642.2010.00725.x

Elsayed, O.R., ElKot, Y.G., ElRefaai, D.A., Abdelfattah, H.M., ElSayed, M., Hamdy, A., 2023. Automated Identification and Classification of Teeth Fossils, in: 2023 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC). Presented at the 2023 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC), IEEE, Cairo, Egypt, pp. 179–186. https://doi.org/10.1109/MIUCC58832.2023.10278368

Eronen, J.T., Polly, P.D., Fred, M., Damuth, J., Frank, D.C., Mosbrugger, V., Scheidegger, C., Stenseth, N.Chr., Fortelius, M., 2010. Ecometrics: The traits that bind the past and present together. Integrative Zoology 5, 88–101. https://doi.org/10.1111/j.1749-4877.2010.00192.x

European Commission, 2018. A definition of AI: Main capabilities and scientific disciplines.

Ezard, T.H., Pearson, P.N., Purvis, A., 2010. Algorithmic approaches to aid species' delimitation in multidimensional morphospace. BMC Evol Biol 10, 175. https://doi.org/10.1186/1471-2148-10-175

Felice, R.N., Watanabe, A., Cuff, A.R., Hanson, M., Bhullar, B.-A.S., Rayfield, E.R., Witmer, L.M., Norell, M.A., Goswami, A., 2020. Decelerated dinosaur skull evolution with the origin of birds. PLoS Biol 18, e3000801. https://doi.org/10.1371/journal.pbio.3000801

Feltes, B.C., Grisci, B.I., Poloni, J.D.F., Dorn, M., 2018. Perspectives and applications of machine learning for evolutionary developmental biology. Molecular Omics 14, 289–306. https://doi.org/10.1039/c8mo00111a

Fenberg, P.B., Self, A., Stewart, J.R., Wilson, R.J., Brooks, S.J., 2016. Exploring the universal ecological responses to climate change in a univoltine butterfly. Journal of Animal Ecology 85, 739–748. https://doi.org/10.1111/1365-2656.12492

Feng, D., De Siqueira, A.D., Yang, S., Tran, T., Bodrito, T., Van Der Walt, S., 2021. machine-shop/mothra: v1.0-rc.2. https://doi.org/10.5281/ZENODO.5551110

Fernandes, A.F.A., Turra, E.M., De Alvarenga, É.R., Passafaro, T.L., Lopes, F.B., Alves, G.F.O., Singh, V., Rosa, G.J.M., 2020. Deep Learning image segmentation for extraction of fish body measurements and prediction of body weight and carcass traits in Nile tilapia. Computers and Electronics in Agriculture 170, 105274. https://doi.org/10.1016/j.compag.2020.105274

Ferreira, A.C., Silva, L.R., Renna, F., Brandl, H.B., Renoult, J.P., Farine, D.R., Covas, R., Doutrelant, C., 2020. Deep learning-based methods for individual recognition in small birds. Methods Ecol Evol 11, 1072–1085. https://doi.org/10.1111/2041-210X.13436

Filella, J.B., Quispe Bonilla, C.C., Quispe, E., Dalerum, F., 2023. Artificial intelligence as a potential tool for micro-histological analysis of herbivore diets. Eur J Wildl Res 69, 11. https://doi.org/10.1007/s10344-022-01640-4

Fishial.ai [WWW Document], 2019. URL https://www.fishial.ai/ (accessed 1.12.24).

Foister, T.I.F., Felice, R.N., 2021. Ecometric modelling of limb proportions and vegetation index among non-human primates in South America. Hystrix It. J. Mamm. 32, 55–59. https://doi.org/10.4404/hystrix-00374-2020

Foote, M., 1997. The Evolution of Morphological Diversity. Annu. Rev. Ecol. Syst. 28, 129–152. https://doi.org/10.1146/annurev.ecolsys.28.1.129

Foote, M., 1993. Discordance and concordance between morphological and taxonomic diversity. Paleobiology 19, 185–204. https://doi.org/10.1017/S0094837300015864

Fraley, C., Raftery, A.E., 2002. Model-Based Clustering, Discriminant Analysis, and Density Estimation. Journal of the American Statistical Association 97, 611–631. https://doi.org/10.1198/016214502760047131

Freitas, M.V., Lemos, C.G., Ariede, R.B., Agudelo, J.F.G., Neto, R.R.O., Borges, C.H.S., Mastrochirico-Filho, V.A., Porto-Foresti, F., Iope, R.L., Batista, F.M., Brega, J.R.F., Hashimoto, D.T., 2023. High-throughput phenotyping by deep learning to include body shape in the breeding program of pacu (Piaractus mesopotamicus). Aquaculture 562, 738847. https://doi.org/10.1016/j.aquaculture.2022.738847

Frey, L., 2018. Artificial Intelligence and Integrated Genotype–Phenotype Identification. Genes 10, 18. https://doi.org/10.3390/genes10010018

Fukushima, K., 1980. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biol. Cybernetics 36, 193–202. https://doi.org/10.1007/BF00344251

Furat, O., Wang, M., Neumann, M., Petrich, L., Weber, M., Krill, C.E., Schmidt, V., 2019. Machine Learning Techniques for the Segmentation of Tomographic Image Data of Functional Materials. Front. Mater. 6, 145. https://doi.org/10.3389/fmats.2019.00145

Galbusera, F., Cina, A., Panico, M., Albano, D., Messina, C., 2020. Image-based biomechanical models of the musculoskeletal system. European Radiology Experimental 4, 49. https://doi.org/10.1186/s41747-020-00172-3

Garcia-Garcia, A., Orts-Escolano, S., Oprea, S., Villena-Martinez, V., Martinez-Gonzalez, P., Garcia-Rodriguez, J., 2018. A survey on deep learning techniques for image and video semantic segmentation. Applied Soft Computing 70, 41–65. https://doi.org/10.1016/j.asoc.2018.05.018

Gardiner, J.D., Behnsen, J., Brassey, C.A., 2018. Alpha shapes: determining 3D shape complexity across morphologically diverse structures. BMC Evol Biol 18, 184. https://doi.org/10.1186/s12862-018-1305-z

Geng, C., Huang, S.-J., Chen, S., 2021. Recent Advances in Open Set Recognition: A Survey. IEEE Trans. Pattern Anal. Mach. Intell. 43, 3614–3631. https://doi.org/10.1109/TPAMI.2020.2981604

Goëau, H., Bonnet, P., Joly, A., Bakić, V., Barbe, J., Yahiaoui, I., Selmi, S., Carré, J., Barthélémy, D., Boujemaa, N., Molino, J.-F., Duché, G., Péronnet, A., 2013. Pl@ntNet mobile app, in: Proceedings of the 21st ACM International Conference on Multimedia. Presented at the MM '13: ACM Multimedia Conference, ACM, Barcelona Spain, pp. 423–424. https://doi.org/10.1145/2502081.2502251

Goëau, H., Lorieul, T., Heuret, P., Joly, A., Bonnet, P., 2022. Can Artificial Intelligence Help in the Study of Vegetative Growth Patterns from Herbarium Collections? An Evaluation of the Tropical Flora of the French Guiana Forest. Plants 11. https://doi.org/10.3390/plants11040530

Goodfellow, I.J., Bengio, Y., Courville, A., 2016. Deep Learning. MIT Press, Cambridge, MA,

USA.
Goswami, A., 2015. Phenome10K: a free online repository for 3-D scans of biological and palaeontological specimens.
Goswami, A., Noirault, E., Coombs, E.J., Clavel, J., Fabre, A.-C., Halliday, T.J.D., Churchill, M., Curtis, A., Watanabe, A., Simmons, N.B., Beatty, B.L., Geisler, J.H., Fox, D.L., Felice, R.N., 2022. Attenuated evolution of mammals through the Cenozoic. Science 378, 377–383. https://doi.org/10.1126/science.abm7525
Goswami, A., Clavel, J., 2024. Morphological evolution in a time of Phenomics (preprint). Life Sciences. https://doi.org/10.32942/X22G7Q
Graves, A., Mohamed, A., Hinton, G., 2013. Speech recognition with deep recurrent neural networks, in: 2013 IEEE International Conference on Acoustics, Speech and Signal Processing. pp. 6645–6649. https://doi.org/10.1109/ICASSP.2013.6638947
Grys, B.T., Lo, D.S., Sahin, N., Kraus, O.Z., Morris, Q., Boone, C., Andrews, B.J., 2017. Machine learning and computer vision approaches for phenotypic profiling. Journal of Cell Biology 216, 65–71. https://doi.org/10.1083/jcb.201610026
Gu, Q., Scott, B., Smith, V., 2023. Planetary Knowledge Base: Semantic Transcription Using Graph Neural Networks. BISS 7, e111168. https://doi.org/10.3897/biss.7.111168
Gu, Q., Scott, B., Smith, V., 2022. Enhancing Botanical Knowledge Graphs with Machine Learning. BISS 6, e91384. https://doi.org/10.3897/biss.6.91384
Guo, S., Xu, P., Miao, Q., Shao, G., Chapman, C.A., Chen, X., He, G., Fang, D., Zhang, H., Sun, Y., Shi, Z., Li, B., 2020. Automatic Identification of Individual Primates with Deep Learning Techniques. iScience 23, 101412. https://doi.org/10.1016/j.isci.2020.101412
Haghighat, M., Browning, L., Sirinukunwattana, K., Malacrino, S., Khalid Alham, N., Colling, R., Cui, Y., Rakha, E., Hamdy, F.C., Verrill, C., Rittscher, J., 2022. Automated quality assessment of large digitised histology cohorts by artificial intelligence. Sci Rep 12, 5002. https://doi.org/10.1038/s41598-022-08351-5
Hallou, A., Yevick, H.G., Dumitrascu, B., Uhlmann, V., 2021. Deep learning for bioimage analysis in developmental biology. Development 148, dev199616. https://doi.org/10.1242/dev.199616
Hanocka, R., Hertz, A., Fish, N., Giryes, R., Fleishman, S., Cohen-Or, D., 2019. MeshCNN: a network with an edge. ACM Trans. Graph. 38, 1–12. https://doi.org/10.1145/3306346.3322959
Hansen, O.L.P., Svenning, J., Olsen, K., Dupont, S., Garner, B.H., Iosifidis, A., Price, B.W., Høye, T.T., 2020. Species-level image classification with convolutional neural network enables insect identification from habitus images. Ecology and Evolution 10, 737–747. https://doi.org/10.1002/ece3.5921
Hartman, E., Sukurdeep, Y., Charon, N., Klassen, E., Bauer, M., 2021. Supervised Deep Learning of Elastic SRV Distances on the Shape Space of Curves, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. pp. 4425–4433.
Hartman, E., Sukurdeep, Y., Klassen, E., Charon, N., Bauer, M., 2023. Elastic Shape Analysis of Surfaces with Second-Order Sobolev Metrics: A Comprehensive Numerical Framework. Int J Comput Vis 131, 1183–1209. https://doi.org/10.1007/s11263-022-01743-0
He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R., 2021. Masked Autoencoders Are Scalable Vision Learners.
He, K., Fan, H., Wu, Y., Xie, S., Girshick, R., 2020. Momentum Contrast for Unsupervised Visual Representation Learning.
He, Y., Cooney, C.R., Maddock, S., Thomas, G.H., 2023. Using pose estimation to identify regions and points on natural history specimens. PLOS Computational Biology 19,

e1010933. https://doi.org/10.1371/journal.pcbi.1010933

He, Y., Varley, Z.K., Nouri, L.O., Moody, C.J.A., Jardine, M.D., Maddock, S., Thomas, G.H., Cooney, C.R., 2022. Deep learning image segmentation reveals patterns of UV reflectance evolution in passerine birds. Nature Communications 13, 5068. https://doi.org/10.1038/s41467-022-32586-5

Hedrick, B.P., Heberling, J.M., Meineke, E.K., Turner, K.G., Grassa, C.J., Park, D.S., Kennedy, J., Clarke, J.A., Cook, J.A., Blackburn, D.C., Edwards, S.V., Davis, C.C., 2020. Digitization and the Future of Natural History Collections. BioScience 70, 243–251. https://doi.org/10.1093/biosci/biz163

Hennig, W., 1966. Phylogenetic systematics. University of Illinois Press, Urbana Chicago London.

Herbst, E.C., Meade, L.E., Lautenschlager, S., Fioritti, N., Scheyer, T.M., 2022. A toolbox for the retrodeformation and muscle reconstruction of fossil specimens in Blender. Royal Society Open Science 9, 220519. https://doi.org/10.1098/rsos.220519

Ho, L.S.T., Dinh, V., Nguyen, C.V., 2019. Multi-task learning improves ancestral state reconstruction. Theoretical Population Biology 126, 33–39. https://doi.org/10.1016/j.tpb.2019.01.001

Hochreiter, S., Schmidhuber, J., 1996. LSTM can Solve Hard Long Time Lag Problems, in: Mozer, M.C., Jordan, M., Petsche, T. (Eds.), Advances in Neural Information Processing Systems. MIT Press.

Hollister, J.D., Cai, X., Horton, T., Price, B.W., Zarzyczny, K.M., Fenberg, P.B., 2023. Using computer vision to identify limpets from their shells: a case study using four species from the Baja California peninsula. Front. Mar. Sci. 10, 1167818. https://doi.org/10.3389/fmars.2023.1167818

Holm, E.A., Cohn, R., Gao, N., Kitahara, A.R., Matson, T.P., Lei, B., Yarasi, S.R., 2020. Overview: Computer Vision and Machine Learning for Microstructural Characterization and Analysis. Metall Mater Trans A 51, 5985–5999. https://doi.org/10.1007/s11661-020-06008-4

Hou, J., He, Y., Yang, H., Connor, T., Gao, J., Wang, Y., Zeng, Y., Zhang, J., Huang, J., Zheng, B., Zhou, S., 2020. Identification of animal individuals using deep learning: A case study of giant panda. Biological Conservation 242, 108414. https://doi.org/10.1016/j.biocon.2020.108414

Hou, Y., Canul-Ku, M., Cui, X., Hasimoto-Beltran, R., Zhu, M., 2021. Semantic segmentation of vertebrate microfossils from computed tomography data using a deep learning approach. J. Micropalaeontol. 40, 163–173. https://doi.org/10.5194/jm-40-163-2021

Hsiang, A.Y., Brombacher, A., Rillo, M.C., Mleneck-Vautravers, M.J., Conn, S., Lordsmith, S., Jentzen, A., Henehan, M.J., Metcalfe, B., Fenton, I.S., Wade, B.S., Fox, L., Meilland, J., Davis, C.V., Baranowski, U., Groeneveld, J., Edgar, K.M., Movellan, A., Aze, T., Dowsett, H.J., Miller, C.G., Rios, N., Hull, P.M., 2019. Endless Forams: >34,000 Modern Planktonic Foraminiferal Images for Taxonomic Training and Automated Species Recognition Using Convolutional Neural Networks. Paleoceanog and Paleoclimatol 34, 1157–1177. https://doi.org/10.1029/2019PA003612

Hu, W., Zhang, C., Jiang, Y., Huang, C., Liu, Q., Xiong, L., Yang, W., Chen, F., 2020. Nondestructive 3D Image Analysis Pipeline to Extract Rice Grain Traits Using X-Ray Computed Tomography. Plant Phenomics 2020, 2020/3414926. https://doi.org/10.34133/2020/3414926

Huang, T., Huang, Y., Lin, W., 2013. Real-time horse gait synthesis. Computer Animation & Virtual 24, 87–95. https://doi.org/10.1002/cav.1469

Hudson, L.N., Blagoderov, V., Heaton, A., Holtzhausen, P., Livermore, L., Price, B.W., Van Der Walt, S., Smith, V.S., 2015. Inselect: Automating the Digitization of Natural History

Collections. PLoS ONE 10, e0143402. https://doi.org/10.1371/journal.pone.0143402

Hughes, E.C., Edwards, D.P., Bright, J.A., Capp, E.J.R., Cooney, C.R., Varley, Z.K., Thomas, G.H., 2022. Global biogeographic patterns of avian morphological diversity. Ecology Letters 25, 598–610. https://doi.org/10.1111/ele.13905

Huiskes, R., Hollister, S.J., 1993. From Structure to Process, From Organ to Cell: Recent Developments of FE-Analysis in Orthopaedic Biomechanics. Journal of Biomechanical Engineering 115, 520–527. https://doi.org/10.1115/1.2895534

Hussein, B.R., Malik, O.A., Ong, W.H., Slik, J.W.F., 2021. Automated extraction of phenotypic leaf traits of individual intact herbarium leaves from herbarium specimen images using deep learning based semantic segmentation. Sensors 21. https://doi.org/10.3390/s21134549

Islam, T., Kim, C.H., Iwata, H., Shimono, H., Kimura, A., 2023. DeepCGP: A Deep Learning Method to Compress Genome-Wide Polymorphisms for Predicting Phenotype of Rice. IEEE/ACM Trans. Comput. Biol. and Bioinf. 20, 2078–2088. https://doi.org/10.1109/TCBB.2022.3231466

Jeanray, N., Marée, R., Pruvot, B., Stern, O., Geurts, P., Wehenkel, L., Muller, M., 2015. Phenotype classification of zebrafish embryos by supervised learning. PLoS ONE 10, 1–20. https://doi.org/10.1371/journal.pone.0116989

Jermyn, I.H., Kurtek, S., Laga, H., Srivastava, A., 2017. Elastic Shape Analysis of Three-Dimensional Objects, Synthesis Lectures on Computer Vision. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-031-01819-0

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: Convolutional Architecture for Fast Feature Embedding.

Jin, B., Cruz, L., Goncalves, N., 2022. Pseudo RGB-D Face Recognition. IEEE Sensors J. 22, 21780–21794. https://doi.org/10.1109/JSEN.2022.3197235

Jolliffe, I.T., Cadima, J., 2016. Principal component analysis: a review and recent developments. Phil. Trans. R. Soc. A. 374, 20150202. https://doi.org/10.1098/rsta.2015.0202

Jones, K.E., Dickson, B.V., Angielczyk, K.D., Pierce, S.E., 2021. Adaptive landscapes challenge the “lateral-to-sagittal” paradigm for mammalian vertebral evolution. Current Biology 31, 1883–1892.

Kale, R.S., Thorat, S., 2021. Image Segmentation Techniques with Machine Learning. IJSRCSEIT 232–235. https://doi.org/10.32628/CSEIT1217653

Kammerer, C.F., Deutsch, M., Lungmus, J.K., Angielczyk, K.D., 2020. Effects of taphonomic deformation on geometric morphometric analysis of fossils: a study using the dicynodont *Diictodon feliceps* (Therapsida, Anomodontia). PeerJ 8, e9925. https://doi.org/10.7717/peerj.9925

Karnani, K., Pepper, J., Bakiş, Y., Wang, X., Bart, H., Breen, D.E., Greenberg, J., 2022. Computational metadata generation methods for biological specimen image collections. International Journal on Digital Libraries. https://doi.org/10.1007/s00799-022-00342-1

Karuppaiah, V., Maruthadurai, R., Das, B., Soumia, P.S., Gadge, A.S., Thangasamy, A., Ramesh, S.V., Shirsat, D.V., Mahajan, V., Krishna, H., Singh, M., 2023. Predicting the potential geographical distribution of onion thrips, Thrips tabaci in India based on climate change projections using MaxEnt. Sci Rep 13, 7934. https://doi.org/10.1038/s41598-023-35012-y

Kass, M., Witkin, A., Terzopoulos, D., 1988. Snakes: Active contour models. Int J Comput Vision 1, 321–331. https://doi.org/10.1007/BF00133570

Kavzoglu, T., 2009. Increasing the accuracy of neural network classification using refined training data. Environmental Modelling & Software 24, 850–858. https://doi.org/10.1016/j.envsoft.2008.11.012

Kendrick, C., Buckley, M., Brassey, C., 2022. MiTiSegmenter: Software for high throughput segmentation and meshing of microCT data in microtiter plate arrays. MethodsX 9, 101849. https://doi.org/10.1016/j.mex.2022.101849

Kiel, S., 2021. Assessing bivalve phylogeny using Deep Learning and Computer Vision approaches (preprint). Evolutionary Biology. https://doi.org/10.1101/2021.04.08.438943

Kikinis, R., Pieper, S.D., Vosburgh, K.G., 2013. 3D Slicer: a platform for subject-specific image analysis, visualization, and clinical support, in: Intraoperative Imaging and Image-Guided Therapy. Springer, pp. 277–289.

King, B., Lee, M.S.Y., 2015. Ancestral State Reconstruction, Rate Heterogeneity, and the Evolution of Reptile Viviparity. Systematic Biology 64, 532–544. https://doi.org/10.1093/sysbio/syv005

Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.-Y., Dollár, P., Girshick, R., 2023. Segment Anything. https://doi.org/10.48550/arXiv.2304.02643

Kishor Kumar, Senthil Kumar, R., Sankar, V., Sakthivel, T., Karunakaran, G., Tripathi, P., 2017. Non-destructive estimation of leaf area of durian (*Durio zibethinus*) – An artificial neural network approach. Scientia Horticulturae 219, 319–325. https://doi.org/10.1016/j.scienta.2017.03.028

Klassen, E., Srivastava, A., 2006. Geodesics Between 3D Closed Curves Using Path-Straightening, in: Leonardis, A., Bischof, H., Pinz, A. (Eds.), Computer Vision – ECCV 2006. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 95–106. https://doi.org/10.1007/11744023_8

Kong, X., Li, J., 2018. Vision-Based Fatigue Crack Detection of Steel Structures Using Video Feature Tracking. Computer-Aided Civil and Infrastructure Engineering 33, 783–799. https://doi.org/10.1111/mice.12353

Korfmann, K., Gaggiotti, O.E., Fumagalli, M., 2023. Deep Learning in Population Genetics. Genome Biology and Evolution 15, evad008. https://doi.org/10.1093/gbe/evad008

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems 25.

Kuhn, T., Hettich, J., Davtyan, R., Gebhardt, J.C.M., 2021. Single molecule tracking and analysis framework including theory-predicted parameter settings. Sci Rep 11, 9465. https://doi.org/10.1038/s41598-021-88802-7

Kwon, Y., Kang, S., Choi, Y.-S., Kim, I., 2021. Evolutionary design of molecules based on deep learning and a genetic algorithm. Sci Rep 11, 17304. https://doi.org/10.1038/s41598-021-96812-8

Lacoste, A., Luccioni, A., Schmidt, V., Dandres, T., 2019. Quantifying the Carbon Emissions of Machine Learning.

Lahiri, S., Robinson, D., Klassen, E., 2015. Precise Matching of PL Curves in $R^N$ in the Square Root Velocity Framework. https://doi.org/10.48550/arXiv.1501.00577

Lannelongue, L., Aronson, H.-E.G., Bateman, A., Birney, E., Caplan, T., Juckes, M., McEntyre, J., Morris, A.D., Reilly, G., Inouye, M., 2023. GREENER principles for environmentally sustainable computational science. Nat Comput Sci 3, 514–521. https://doi.org/10.1038/s43588-023-00461-y

Lannelongue, L., Grealey, J., Inouye, M., 2021. Green Algorithms: Quantifying the Carbon Footprint of Computation. Advanced Science 8, 2100707. https://doi.org/10.1002/advs.202100707

Lautenschlager, S., 2016. Reconstructing the past: methods and techniques for the digital restoration of fossils. Royal Society Open Science 3, 160342. https://doi.org/10.1098/rsos.160342

Le Guillarme, N., Thuiller, W., 2022. TaxoNERD: Deep neural models for the recognition of taxonomic entities in the ecological and evolutionary literature. Methods Ecol Evol 13, 625–641. https://doi.org/10.1111/2041-210X.13778

LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. Nature 521, 436–444. https://doi.org/10.1038/nature14539

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W., Jackel, L.D., 1989. Backpropagation Applied to Handwritten Zip Code Recognition. Neural Computation 1, 541–551. https://doi.org/10.1162/neco.1989.1.4.541

Lee, M.S.Y., Palci, A., 2015. Morphological Phylogenetics in the Genomic Age. Current Biology 25, R922–R929. https://doi.org/10.1016/j.cub.2015.07.009

Lenton, K.A., Nacamuli, R.P., Wan, D.C., Helms, J.A., Longaker, M.T., 2005. Cranial Suture Biology, in: Current Topics in Developmental Biology. Academic Press, pp. 287–328. https://doi.org/10.1016/S0070-2153(05)66009-7

Li, X., Zhang, Y., Wu, J., Dai, Q., 2023. Challenges and opportunities in bioimage analysis. Nat Methods 20, 958–961. https://doi.org/10.1038/s41592-023-01900-4

Lin, D., Dai, J., Jia, J., He, K., Sun, J., 2016. ScribbleSup: Scribble-Supervised Convolutional Networks for Semantic Segmentation, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Liu, G.R., 2019. FEA-AI and AI-AI: Two-Way Deepnets for Real-Time Computations for Both Forward and Inverse Mechanics Problems. Int. J. Comput. Methods 16, 1950045. https://doi.org/10.1142/S0219876219500452

Liu, Y., Wang, S.-L., Zhang, J.-F., Zhang, W., Zhou, S., Li, W., 2021. DMFMDA: Prediction of Microbe-Disease Associations Based on Deep Matrix Factorization Using Bayesian Personalized Ranking. IEEE/ACM Trans. Comput. Biol. and Bioinf. 18, 1763–1772. https://doi.org/10.1109/TCBB.2020.3018138

Liu, Zhichao, Jin, L., Chen, J., Fang, Q., Ablameyko, S., Yin, Z., Xu, Y., 2021. A survey on applications of deep learning in microscopy image analysis. Computers in Biology and Medicine 134, 104523. https://doi.org/10.1016/j.compbiomed.2021.104523

Liu, Ze, Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B., 2021. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows.

Lloyd, S., 1982. Least squares quantization in PCM. IEEE Transactions on Information Theory 28, 129–137. https://doi.org/10.1109/TIT.1982.1056489

Long, E., Wan, P., Chen, Q., Lu, Z., Choi, J., 2023. From function to translation: Decoding genetic susceptibility to human diseases via artificial intelligence. Cell Genomics 3, 100320. https://doi.org/10.1016/j.xgen.2023.100320

Lösel, P., Heuveline, V., 2017. A GPU Based Diffusion Method for Whole-Heart and Great Vessel Segmentation, in: Zuluaga, M.A., Bhatia, K., Kainz, B., Moghari, M.H., Pace, D.F. (Eds.), Reconstruction, Segmentation, and Analysis of Medical Images. Springer International Publishing, Cham, pp. 121–128. https://doi.org/10.1007/978-3-319-52280-7_12

Lösel, P.D., Monchanin, C., Lebrun, R., Jayme, A., Relle, J.J., Devaud, J.-M., Heuveline, V., Lihoreau, M., 2023. Natural variability in bee brain size and symmetry revealed by micro-CT imaging and deep learning. PLoS Comput Biol 19, e1011529. https://doi.org/10.1371/journal.pcbi.1011529

Lösel, P.D., Van De Kamp, T., Jayme, A., Ershov, A., Faragó, T., Pichler, O., Tan Jerome, N., Aadepu, N., Bremer, S., Chilingaryan, S.A., Heethoff, M., Kopmann, A., Odar, J., Schmelzle, S., Zuber, M., Wittbrodt, J., Baumbach, T., Heuveline, V., 2020. Introducing Biomedisa as an open-source online platform for biomedical image segmentation. Nat Commun 11, 5577. https://doi.org/10.1038/s41467-020-19303-w

Love, N.L.R., Bonnet, P., Goëau, H., Joly, A., Mazer, S.J., 2021. Machine learning undercounts reproductive organs on herbarium specimens but accurately derives their quantitative phenological status: A case study of streptanthus tortuosus. Plants 10.

https://doi.org/10.3390/plants10112471

Low, B.W., Zeng, Y., Tan, H.H., Yeo, D.C.J., 2021. Predictor complexity and feature selection affect Maxent model transferability: Evidence from global freshwater invasive species. Diversity and Distributions 27, 497–511. https://doi.org/10.1111/ddi.13211

Lowe, D.G., 2004. Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision 60, 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Lu, Y., Wang, R., Hu, T., He, Q., Chen, Z.S., Wang, J., Liu, L., Fang, C., Luo, J., Fu, L., Yu, L., Liu, Q., 2023. Nondestructive 3D phenotyping method of passion fruit based on X-ray micro-computed tomography and deep learning. Front. Plant Sci. 13, 1087904. https://doi.org/10.3389/fpls.2022.1087904

Luo, D., Zeng, W., Chen, J., Tang, W., 2021. Deep Learning for Automatic Image Segmentation in Stomatology and Its Clinical Application. Front. Med. Technol. 3, 767836. https://doi.org/10.3389/fmedt.2021.767836

Lürig, M.D., Donoughe, S., Svensson, E.I., Porto, A., Tsuboi, M., 2021. Computer Vision, Machine Learning, and the Promise of Phenomics in Ecology and Evolutionary Biology. Front. Ecol. Evol. 9, 642774. https://doi.org/10.3389/fevo.2021.642774

Ma, Y., Zhi, Z., 2022. Extraction and Analysis of Foot Bone Shape Features Based on Deep Learning. Computational and Mathematical Methods in Medicine 2022, 1–11. https://doi.org/10.1155/2022/2372160

MacLeod, N., Price, B., Stevens, Z., 2022. What you sample is what you get: ecomorphological variation in Trithemis (Odonata, Libellulidae) dragonfly wings reconsidered. BMC Ecol Evo 22, 43. https://doi.org/10.1186/s12862-022-01978-y

Maddison, W.P., Knowles, L.L., 2006. Inferring Phylogeny Despite Incomplete Lineage Sorting. Systematic Biology 55, 21–30. https://doi.org/10.1080/10635150500354928

Mäder, P., Boho, D., Rzanny, M., Seeland, M., Wittich, H.C., Deggelmann, A., Wäldchen, J., 2021. The Flora Incognita app – Interactive plant species identification. Methods Ecol Evol 12, 1335–1342. https://doi.org/10.1111/2041-210X.13611

Maga, A.M., Tustison, N.J., Avants, B.B., 2017. A population level atlas of *Mus musculus* craniofacial skeleton and automated image-based shape analysis. Journal of Anatomy 231, 433–443. https://doi.org/10.1111/joa.12645

Mahendiran, M., Parthiban, M., Azeez, P.A., 2022. Signals of local bioclimate-driven ecomorphological changes in wild birds. Sci Rep 12, 15815. https://doi.org/10.1038/s41598-022-20041-w

Maia, R., Gruson, H., Endler, J.A., White, T.E., 2019. pavo 2: New tools for the spectral and spatial analysis of colour in r. Methods Ecol Evol 10, 1097–1107. https://doi.org/10.1111/2041-210X.13174

Marks, M., Jin, Q., Sturman, O., Von Ziegler, L., Kollmorgen, S., Von Der Behrens, W., Mante, V., Bohacek, J., Yanik, M.F., 2022. Deep-learning-based identification, tracking, pose estimation and behaviour classification of interacting primates and mice in complex environments. Nat Mach Intell 4, 331–340. https://doi.org/10.1038/s42256-022-00477-5

Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Jia, Y., Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, Xiaoqiang Zheng, 2015. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems.

Martin-Brualla, R., Radwan, N., Sajjadi, M.S., Barron, J.T., Dosovitskiy, A., Duckworth, D., 2021. Nerf in the wild: Neural radiance fields for unconstrained photo collections. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition 7210–7219.

Masaeli, M., Gupta, D., O'Byrne, S., Tse, H.T.K., Gossett, D.R., Tseng, P., Utada, A.S., Jung, H.J., Young, S., Clark, A.T., Di Carlo, D., 2016. Multiparameter mechanical and morphometric screening of cells. Scientific Reports 6, 1–11. https://doi.org/10.1038/srep37863

Mathis, A., Mamidanna, P., Cury, K.M., Abe, T., Murthy, V.N., Mathis, M.W., Bethge, M., 2018. DeepLabCut: markerless pose estimation of user-defined body parts with deep learning. Nat Neurosci 21, 1281–1289. https://doi.org/10.1038/s41593-018-0209-y

McCulloch, W.S., Pitts, W., 1943. A logical calculus of the ideas immanent in nervous activity. Bulletin of Mathematical Biophysics 5, 115–133. https://doi.org/10.1007/BF02478259

McGhee, G.R., 1999. Theoretical morphology: the concept and its applications. Columbia University Press.

McGhee, G.R., 1980. Shell form in the biconvex articulate Brachiopoda: a geometric analysis. Paleobiology 6, 57–76. https://doi.org/10.1017/S0094837300012513

Medina, J.J., Maley, J.M., Sannapareddy, S., Medina, N.N., Gilman, C.M., McCormack, J.E., 2020. A rapid and cost-effective pipeline for digitization of museum specimens with 3D photogrammetry. PLoS ONE 15, e0236417. https://doi.org/10.1371/journal.pone.0236417

Mehrabi, N., Morstatter, F., Saxena, N., Lerman, K., Galstyan, A., 2021. A survey on bias and fairness in machine learning. ACM computing surveys (CSUR) 54, 1–35.

Merow, C., Smith, M.J., Silander, J.A., 2013. A practical guide to MaxEnt for modeling species' distributions: what it does, and why inputs and settings matter. Ecography 36, 1058–1069. https://doi.org/10.1111/j.1600-0587.2013.07872.x

Mieth, B., Rozier, A., Rodriguez, J.A., Höhne, M.M.C., Görnitz, N., Müller, K.-R., 2021. DeepCOMBI: explainable artificial intelligence for the analysis and discovery in genome-wide association studies. NAR Genomics and Bioinformatics 3, lqab065. https://doi.org/10.1093/nargab/lqab065

Milletari, F., Navab, N., Ahmadi, S.-A., 2016. V-Net: Fully Convolutional Neural Networks for Volumetric Medical Image Segmentation, in: 2016 Fourth International Conference on 3D Vision (3DV). Presented at the 2016 Fourth International Conference on 3D Vision (3DV), IEEE, Stanford, CA, USA, pp. 565–571. https://doi.org/10.1109/3DV.2016.79

Minaee, S., Wang, Y., 2019. An ADMM Approach to Masked Signal Decomposition Using Subspace Representation. IEEE Trans. on Image Process. 28, 3192–3204. https://doi.org/10.1109/TIP.2019.2894966

Mitteroecker, P., Schaefer, K., 2022. Thirty years of geometric morphometrics: Achievements, challenges, and the ongoing quest for biological meaningfulness. American Journal of Biological Anthropology 178, 181–210. https://doi.org/10.1002/ajpa.24531

Mo, Y., Hahn, M., Smith, M., 2023. Applications of Machine Learning in Phylogenetics (preprint). Biology. https://doi.org/10.32942/X2XG7G

Moen, E., Bannon, D., Kudo, T., Graf, W., Covert, M., Van Valen, D., 2019. Deep learning for cellular image analysis. Nat Methods 16, 1233–1246. https://doi.org/10.1038/s41592-019-0403-1

Mohammadi, V., Minaei, S., Mahdavian, A.R., Khoshtaghaza, M.H., Gouton, P., 2021. Estimation of Leaf Area in Bell Pepper Plant using Image Processing techniques and Artificial Neural Networks, in: 2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). Presented at the 2021 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), IEEE, Kuala Terengganu, Malaysia, pp. 173–178. https://doi.org/10.1109/ICSIPA52582.2021.9576778

Moi, D., Dessimoz, C., 2022. Reconstructing protein interactions across time using phylogeny-aware graph neural networks (preprint). Bioinformatics. https://doi.org/10.1101/2022.07.21.501014

Mongia, A., Majumdar, A., 2020. Drug-target interaction prediction using Multi Graph Regularized Nuclear Norm Minimization. PLoS ONE 15, e0226484. https://doi.org/10.1371/journal.pone.0226484

Mouloodi, S., Rahmanpanah, H., Gohari, S., Burvill, C., Tse, K.M., Davies, H.M.S., 2021. What can artificial intelligence and machine learning tell us? A review of applications to equine biomechanical research. Journal of the Mechanical Behavior of Biomedical Materials 123, 104728. https://doi.org/10.1016/j.jmbbm.2021.104728

Mulqueeney, J.M., Searle-Barnes, A., Brombacher, A., Sweeney, M., Goswami, A., Ezard, T.H.G., 2024. How many specimens make a sufficient training set for automated 3D feature extraction? https://doi.org/10.1101/2024.01.10.575054

Naert, T., Çiçek, Ö., Ogar, P., Bürgi, M., Shaidani, N.I., Kaminski, M.M., Xu, Y., Grand, K., Vujanovic, M., Prata, D., Hildebrandt, F., Brox, T., Ronneberger, O., Voigt, F.F., Helmchen, F., Loffing, J., Horb, M.E., Willsey, H.R., Lienkamp, S.S., 2021. Deep learning is widely applicable to phenotyping embryonic development and disease. Development (Cambridge) 148, 1–18. https://doi.org/10.1242/dev.199664

Najman, L., Schmitt, M., 1994. Watershed of a continuous function. Signal Processing 38, 99–112. https://doi.org/10.1016/0165-1684(94)90059-0

Narayana, P.A., Coronado, I., Sujit, S.J., Wolinsky, J.S., Lublin, F.D., Gabr, R.E., 2020. Deep-Learning-Based Neural Tissue Segmentation of MRI in Multiple Sclerosis: Effect of Training Set Size. Magnetic Resonance Imaging 51, 1487–1496. https://doi.org/10.1002/jmri.26959

Nath, T., Mathis, A., Chen, A.C., Patel, A., Bethge, M., Mathis, M.W., 2019. Using DeepLabCut for 3D markerless pose estimation across species and behaviors. Nat Protoc 14, 2152–2176. https://doi.org/10.1038/s41596-019-0176-0

Nesterenko, L., Boussau, B., Jacob, L., 2022. Phyloformer: towards fast and accurate phylogeny estimation with self-attention networks (preprint). Evolutionary Biology. https://doi.org/10.1101/2022.06.24.496975

Nikolados, E.-M., Wongprommoon, A., Aodha, O.M., Cambray, G., Oyarzún, D.A., 2022. Accuracy and data efficiency in deep learning models of protein expression. Nat Commun 13, 7755. https://doi.org/10.1038/s41467-022-34902-5

Nock, R., Nielsen, F., 2004. Statistical region merging. IEEE Trans. Pattern Anal. Machine Intell. 26, 1452–1458. https://doi.org/10.1109/TPAMI.2004.110

Norouzzadeh, M.S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M.S., Packer, C., Clune, J., 2018. Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. Proc. Natl. Acad. Sci. U.S.A. 115. https://doi.org/10.1073/pnas.1719367115

Oord, A. van den, Li, Y., Vinyals, O., 2019. Representation Learning with Contrastive Predictive Coding.

Osher, S., Sethian, J.A., 1988. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. Journal of Computational Physics 79, 12–49. https://doi.org/10.1016/0021-9991(88)90002-2

Otsu, N., 1979. A Threshold Selection Method from Gray-Level Histograms. IEEE Trans. Syst., Man, Cybern. 9, 62–66. https://doi.org/10.1109/TSMC.1979.4310076

Padilla-García, N., Šrámková, G., Záveská, E., Šlenker, M., Clo, J., Zeisek, V., Lučanová, M., Rurane, I., Kolář, F., Marhold, K., 2023. The importance of considering the evolutionary history of polyploids when assessing climatic niche evolution. Journal of Biogeography 50, 86–100. https://doi.org/10.1111/jbi.14496

Papageorgiou, C.P., Oren, M., Poggio, T., 1998. A general framework for object detection, in: Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271). Presented at the Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271), pp. 555–562. https://doi.org/10.1109/ICCV.1998.710772

Parker, A.K., Müller, J., Boisserie, J.-R., Head, J.J., 2023. The utility of body size as a functional trait to link the past and present in a diverse reptile clade. Proc. Natl. Acad. Sci. U.S.A. 120, e2201948119. https://doi.org/10.1073/pnas.2201948119

Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., Killeen, T., Lin, Z., Gimelshein, N., Antiga, L., Desmaison, A., Kopf, A., Yang, E., DeVito, Z., Raison, M., Tejani, A., Chilamkurthy, S., Steiner, B., Fang, L., Bai, J., Chintala, S., 2019. PyTorch: An Imperative Style, High-Performance Deep Learning Library, in: Wallach, H., Larochelle, H., Beygelzimer, A., Alché-Buc, F. d', Fox, E., Garnett, R. (Eds.), Advances in Neural Information Processing Systems 32. Curran Associates, Inc., pp. 8024–8035.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in Python. Journal of Machine Learning Research 12, 2825–2830.

Pereira, T.D., Aldarondo, D.E., Willmore, L., Kislin, M., Wang, S.S.-H., Murthy, M., Shaevitz, J.W., 2019. Fast animal pose estimation using deep neural networks. Nat Methods 16, 117–125. https://doi.org/10.1038/s41592-018-0234-5

Perera, P., Patel, V.M., 2019. Learning Deep Features for One-Class Classification. IEEE Trans. on Image Process. 28, 5450–5463. https://doi.org/10.1109/TIP.2019.2917862

Perronnin, F., Dance, C., 2007. Fisher Kernels on Visual Vocabularies for Image Categorization, in: 2007 IEEE Conference on Computer Vision and Pattern Recognition. Presented at the 2007 IEEE Conference on Computer Vision and Pattern Recognition, IEEE, Minneapolis, MN, USA, pp. 1–8. https://doi.org/10.1109/CVPR.2007.383266

Perronnin, F., Sánchez, J., Mensink, T., 2010. Improving the fisher kernel for large-scale image classification, in: Computer Vision–ECCV 2010: 11th European Conference on Computer Vision, Heraklion, Crete, Greece, September 5-11, 2010, Proceedings, Part IV 11. Springer, pp. 143–156.

Philippe, H., Poustka, A.J., Chiodin, M., Hoff, K.J., Dessimoz, C., Tomiczek, B., Schiffer, P.H., Müller, S., Domman, D., Horn, M., Kuhl, H., Timmermann, B., Satoh, N., Hikosaka-Katayama, T., Nakano, H., Rowe, M.L., Elphick, M.R., Thomas-Chollier, M., Hankeln, T., Mertes, F., Wallberg, A., Rast, J.P., Copley, R.R., Martinez, P., Telford, M.J., 2019. Mitigating Anticipated Effects of Systematic Errors Supports Sister-Group Relationship between Xenacoelomorpha and Ambulacraria. Current Biology 29, 1818-1826.e6. https://doi.org/10.1016/j.cub.2019.04.009

Phillips, S.J., Anderson, R.P., Schapire, R.E., 2006. Maximum entropy modeling of species geographic distributions. Ecological Modelling 190, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Phillips, S.J., Dudík, M., Schapire, R.E., 2024. Maxent software for modeling species niches and distributions.

Pichler, M., Hartig, F., 2023. Machine learning and deep learning—A review for ecologists. Methods Ecol Evol 14, 994–1016. https://doi.org/10.1111/2041-210X.14061

Pierson, E., Daoudi, M., Tumpach, A.-B., 2021. A Riemannian Framework for Analysis of Human Body Surface. https://doi.org/10.48550/arXiv.2108.11449

Pinheiro, D., Santander-Jimenéz, S., Ilic, A., 2022. PhyloMissForest: a random forest framework to construct phylogenetic trees with missing data. BMC Genomics 23, 377. https://doi.org/10.1186/s12864-022-08540-6

Pl@ntNet IPT, 2023.

Pl@ntNet [WWW Document], 2023. . Pl@ntNet. URL https://plantnet.org/en/ (accessed 1.12.24).

Polly, P.D., Stayton, C.T., Dumont, E.R., Pierce, S.E., Rayfield, E.J., Angielczyk, K.D., 2016. Combining geometric morphometrics and finite element analysis with evolutionary modeling: towards a synthesis. Journal of Vertebrate Paleontology 36, e1111225. https://doi.org/10.1080/02724634.2016.1111225

Pomidor, B.J., Makedonska, J., Slice, D.E., 2016. A Landmark-Free Method for Three-Dimensional Shape Analysis. PLOS ONE 11, e0150368. https://doi.org/10.1371/journal.pone.0150368

Poon, S.T.S., Hanna, F.W.F., Lemarchand, F., George, C., Clark, A., Lea, S., Coleman, C., Sollazzo, G., 2023. Detecting adrenal lesions on 3D CT scans using a 2.5D deep learning model (preprint). Radiology and Imaging. https://doi.org/10.1101/2023.02.22.23286184

Porto, A., Rolfe, S., Maga, A.M., 2021. ALPACA: A fast and accurate computer vision approach for automated landmarking of three-dimensional biological structures. Methods Ecol Evol 12, 2129–2144. https://doi.org/10.1111/2041-210X.13689

Porto, A., Voje, K.L., 2020. ML-morph: A fast, accurate and general approach for automated detection and landmarking of biological structures in images. Methods Ecol Evol 11, 500–512. https://doi.org/10.1111/2041-210X.13373

Pratapa, A., Doron, M., Caicedo, J.C., 2021. Image-based cell phenotyping with deep learning. Current Opinion in Chemical Biology, Mechanistic Biology * Machine Learning in Chemical Biology 65, 9–17. https://doi.org/10.1016/j.cbpa.2021.04.001

Price, B.W., Dupont, S., Allan, E.L., Blagoderov, V., Butcher, A.J., Durrant, J., Holtzhausen, P., Kokkini, P., Livermore, L., Hardy, H., Smith, V., 2018. ALICE: Angled Label Image Capture and Extraction for high throughput insect specimen digitisation (preprint). Open Science Framework. https://doi.org/10.31219/osf.io/s2p73

Punyasena, S.W., Tcheng, D.K., Wesseln, C., Mueller, P.G., 2012. Classifying black and white spruce pollen using layered machine learning. New Phytologist 196, 937–944. https://doi.org/10.1111/j.1469-8137.2012.04291.x

Qin, A., Liu, B., Guo, Q., Bussmann, R.W., Ma, F., Jian, Z., Xu, G., Pei, S., 2017. Maxent modeling for predicting impacts of climate change on the potential distribution of Thuja sutchuenensis Franch., an extremely endangered conifer from southwestern China. Global Ecology and Conservation 10, 139–146. https://doi.org/10.1016/j.gecco.2017.02.004

Qin, Y., Havulinna, A.S., Liu, Y., Jousilahti, P., Ritchie, S.C., Tokolyi, A., Sanders, J.G., Valsta, L., Brożyńska, M., Zhu, Q., Tripathi, A., Vázquez-Baeza, Y., Loomba, R., Cheng, S., Jain, M., Niiranen, T., Lahti, L., Knight, R., Salomaa, V., Inouye, M., Méric, G., 2022. Combined effects of host genetics and diet on human gut microbiota and incident disease in a single population cohort. Nat Genet 54, 134–142. https://doi.org/10.1038/s41588-021-00991-z

Qin, Z., Qin, F., Li, Y., Yu, C., 2022. Intelligent Objective Osteon Segmentation Based on Deep Learning. Front. Earth Sci. 10, 783481. https://doi.org/10.3389/feart.2022.783481

Rabinovich, J.E., 2021. Morphology, Life Cycle, Environmental Factors and Fitness – a Machine Learning Analysis in Kissing Bugs (Hemiptera, Reduviidae, Triatominae). Front. Ecol. Evol. 9, 651683. https://doi.org/10.3389/fevo.2021.651683

Radford, A., Kim, J.W., Hallacy, C., Ramesh, A., Goh, G., Agarwal, S., Sastry, G., Askell, A., Mishkin, P., Clark, J., Krueger, G., Sutskever, I., 2021. Learning Transferable Visual Models From Natural Language Supervision, in: Meila, M., Zhang, T. (Eds.), Proceedings of the 38th International Conference on Machine Learning, Proceedings of Machine Learning Research. PMLR, pp. 8748–8763.

Radford, A., Narasimhan, K., Salimans, T., Sutskever, I., others, 2018. Improving language understanding by generative pre-training.

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., Sutskever, I., others, 2019. Language models are unsupervised multitask learners. OpenAI blog 1, 9.

Raissi, M., Perdikaris, P., Karniadakis, G.E., 2019. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. Journal of Computational Physics 378, 686–707. https://doi.org/10.1016/j.jcp.2018.10.045

Rau, C., Marathe, S., Bodey, A.J., Storm, M., Batey, D., Cipiccia, S., Li, P., Ziesche, R., 2021. High-throughput micro and nano-tomography, in: Müller, B., Wang, G. (Eds.), Developments in X-Ray Tomography XIII. Presented at the Developments in X-Ray Tomography XIII, SPIE, San Diego, United States, p. 49. https://doi.org/10.1117/12.2598470

Ravindran, S., 2022. Five ways deep learning has transformed image analysis. Nature 609, 864–866. https://doi.org/10.1038/d41586-022-02964-6

Ren, S., He, K., Girshick, R., Sun, J., 2016. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.

Rezaeitaleshmahalleh, M., Mu, N., Lyu, Z., Zhou, W., Zhang, X., Rasmussen, T.E., McBane, R.D., Jiang, J., 2023. Radiomic-based Textural Analysis of Intraluminal Thrombus in Aortic Abdominal Aneurysms: A Demonstration of Automated Workflow. J. of Cardiovasc. Trans. Res. https://doi.org/10.1007/s12265-023-10404-7

Robillard, A.J., Trizna, M.G., Ruiz-Tafur, M., Dávila Panduro, E.L., de Santana, C.D., White, A.E., Dikow, R.B., Deichmann, J.L., 2023. Application of a deep learning image classifier for identification of Amazonian fishes. Ecology and Evolution 13, 1–9. https://doi.org/10.1002/ece3.9987

Rolfe, S., Pieper, S., Porto, A., Diamond, K., Winchester, J., Shan, S., Kirveslahti, H., Boyer, D., Summers, A., Maga, A.M., 2021. SlicerMorph: An open and extensible platform to retrieve, visualize and analyse 3D morphology. Methods Ecol Evol 12, 1816–1825. https://doi.org/10.1111/2041-210X.13669

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science. Presented at the International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Munich, Germany, pp. 234–241. https://doi.org/10.1007/978-3-319-24574-4_28

Rosenblatt, F., 1958. The perceptron: A probabilistic model for information storage and organization in the brain. Psychological Review 65, 386–408. https://doi.org/10.1037/h0042519

Ross, C.F., 2005. Finite element analysis in vertebrate biomechanics. The Anatomical Record Part A: Discoveries in Molecular, Cellular, and Evolutionary Biology 283A, 253–258. https://doi.org/10.1002/ar.a.20177

Roure, B., Baurain, D., Philippe, H., 2013. Impact of Missing Data on Phylogenies Inferred from Empirical Phylogenomic Data Sets. Molecular Biology and Evolution 30, 197–214. https://doi.org/10.1093/molbev/mss208

Rowe, T., 2002. DigiMorph [WWW Document]. URL https://www.digimorph.org (accessed 1.15.24).

Ruder, S., 2017. An Overview of Multi-Task Learning in Deep Neural Networks. https://doi.org/10.48550/arXiv.1706.05098

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A.C., Fei-Fei, L., 2015. ImageNet Large Scale Visual

Recognition Challenge.

Russel, S., Norvig, P., 2021. Artificial Intelligence: A Modern Approach, Global Edition. Foundations 19, 1–1166.

Salifu, D., Ibrahim, E.A., Tonnang, H.E.Z., 2022. Leveraging machine learning tools and algorithms for analysis of fruit fly morphometrics. Sci Rep 12, 7208. https://doi.org/10.1038/s41598-022-11258-w

Salili-James, A., Mackay, A., Rodriguez-Alvarez, E., Rodriguez-Perez, D., Mannack, T., Rawlings, T.A., Palmer, A.R., Todd, J., Riutta, T.E., Macinnis-Ng, C., Han, Z., Davies, M., Thorpe, Z., Marsland, S., Leroi, A.M., 2022a. Classifying organisms and artefacts by their outline shapes. J. R. Soc. Interface. 19, 20220493. https://doi.org/10.1098/rsif.2022.0493

Salili-James, A., Scott, B., Smith, V., 2022b. ALICE Software: Machine learning &amp; computer vision for automatic label extraction. BISS 6, e91443. https://doi.org/10.3897/biss.6.91443

Samoili, S., Cobo, M.L., Gomez, E., De Prato, G., Martinez-Plumed, F., Delipetrev, B., 2020. AI Watch. Defining Artificial Intelligence. Towards an operational definition and taxonomy of artificial intelligence [WWW Document]. URL https://publications.jrc.ec.europa.eu/repository/handle/JRC118163 (accessed 1.12.24).

Sapoval, N., Aghazadeh, A., Nute, M.G., Antunes, D.A., Balaji, A., Baraniuk, R., Barberan, C.J., Dannenfelser, R., Dun, C., Edrisi, M., Elworth, R.A.L., Kille, B., Kyrillidis, A., Nakhleh, L., Wolfe, C.R., Yan, Z., Yao, V., Treangen, T.J., 2022. Current progress and open challenges for applying deep learning across the biosciences. Nat Commun 13, 1728. https://doi.org/10.1038/s41467-022-29268-7

Saupe, E.E., Farnsworth, A., Lunt, D.J., Sagoo, N., Pham, K.V., Field, D.J., 2019. Climatic shifts drove major contractions in avian latitudinal distributions throughout the Cenozoic. Proc. Natl. Acad. Sci. U.S.A. 116, 12895–12900. https://doi.org/10.1073/pnas.1903866116

Schlager, S., Profico, A., Vincenzo, F.D., Manzi, G., 2018. Retrodeformation of fossil specimens based on 3D bilateral semi-landmarks: Implementation in the R package “Morpho.” PLOS ONE 13, e0194073. https://doi.org/10.1371/journal.pone.0194073

Schmelzle, S., Van De Kamp, T., Heethoff, M., Heuveline, V., Lösel, P., Becker, J., Beckmann, F., Schluenzen, F., Hammel, J.U., Kopmann, A., Mexner, W., Vogelgesang, M., Jerome, N.T., Betz, O., Beutel, R., Wipfler, B., Blanke, A., Harzsch, S., Hörnig, M., Baumbach, T., 2017. The NOVA project: maximizing beam time efficiency through synergistic analyses of SRμCT data, in: Müller, B., Wang, G. (Eds.), Developments in X-Ray Tomography XI. Presented at the Developments in X-Ray Tomography XI, SPIE, San Diego, United States, p. 24. https://doi.org/10.1117/12.2275959

Schmidt, S., Balke, M., Lafogler, S., 2012. DScan – a high-performance digital scanning system for entomological collections. ZK 209, 183–191. https://doi.org/10.3897/zookeys.209.3115

Schneider, C.A., Rasband, W.S., Eliceiri, K.W., 2012. NIH Image to ImageJ: 25 years of image analysis. Nature methods 9, 671–675.

Schneider, L., Niemann, A., Beuing, O., Preim, B., Saalfeld, S., 2021. MedmeshCNN - Enabling meshcnn for medical surface models. Computer Methods and Programs in Biomedicine 210, 106372. https://doi.org/10.1016/j.cmpb.2021.106372

Schneider, S., Greenberg, S., Taylor, G.W., Kremer, S.C., 2020. Three critical factors affecting automated image species recognition performance for camera traps. Ecology and Evolution 10, 3503–3517. https://doi.org/10.1002/ece3.6147

Schuettpelz, E., Frandsen, P.B., Dikow, R.B., Brown, A., Orli, S., Peters, M., Metallo, A., Funk, V.A., Dorr, L.J., 2017. Applications of deep convolutional neural networks to digitized natural history collections. Biodiversity Data Journal 5.

https://doi.org/10.3897/BDJ.5.e21139

Schwartz, S.T., Alfaro, M.E., 2021. *Sashimi* : A toolkit for facilitating high-throughput organismal image segmentation using deep learning. Methods Ecol Evol 12, 2341–2354. https://doi.org/10.1111/2041-210X.13712

Scott, B., Livermore, L., 2021. Extracting Data at Scale: Machine learning at the Natural History Museum. BISS 5, e74031. https://doi.org/10.3897/biss.5.74031

Scott, B., Salili-James, A., Smith, V., 2023. Robot-in-the-loop: Prototyping robotic digitisation at the Natural History Museum. BISS 7, e112947. https://doi.org/10.3897/biss.7.112947

Shallue, C.J., Vanderburg, A., 2018. Identifying Exoplanets with Deep Learning: A Five-planet Resonant Chain around Kepler-80 and an Eighth Planet around Kepler-90. AJ 155, 94. https://doi.org/10.3847/1538-3881/aa9e09

Sharif Razavian, A., Azizpour, H., Sullivan, J., Carlsson, S., 2014. CNN features off-the-shelf: an astounding baseline for recognition, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Presented at the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 806–813.

Shearer, B.M., Cooke, S.B., Halenar, L.B., Reber, S.L., Plummer, J.E., Delson, E., Tallman, M., 2017. Evaluating causes of error in landmark-based data collection using scanners. PLoS ONE 12, e0187452. https://doi.org/10.1371/journal.pone.0187452

Shi, D., Wang, Y., Ai, Z., 2010. Effect of anterior cruciate ligament reconstruction on biomechanical features of knee in level walking: a meta-analysis. Chinese Medical Journal 123, 3137. https://doi.org/10.3760/cma.j.issn.0366-6999.2010.21.034

Shmatko, A., Ghaffari Laleh, N., Gerstung, M., Kather, J.N., 2022. Artificial intelligence in histopathology: enhancing cancer research and clinical oncology. Nat Cancer 3, 1026–1038. https://doi.org/10.1038/s43018-022-00436-4

Shorten, C., Khoshgoftaar, T.M., 2019. A survey on Image Data Augmentation for Deep Learning. J Big Data 6, 60. https://doi.org/10.1186/s40537-019-0197-0

Shu, Z., Yang, S., Wu, H., Xin, S., Pang, C., Kavan, L., Liu, L., 2022. 3D Shape Segmentation Using Soft Density Peak Clustering and Semi-Supervised Learning. Computer-Aided Design 145, 103181. https://doi.org/10.1016/j.cad.2021.103181

Sillero, N., Barbosa, A.M., 2021. Common mistakes in ecological niche models. International Journal of Geographical Information Science 35, 213–226. https://doi.org/10.1080/13658816.2020.1798968

Silver, D., Huang, A., Maddison, C.J., Guez, A., Sifre, L., Van Den Driessche, G., Schrittwieser, J., Antonoglou, I., Panneershelvam, V., Lanctot, M., Dieleman, S., Grewe, D., Nham, J., Kalchbrenner, N., Sutskever, I., Lillicrap, T., Leach, M., Kavukcuoglu, K., Graepel, T., Hassabis, D., 2016. Mastering the game of Go with deep neural networks and tree search. Nature 529, 484–489. https://doi.org/10.1038/nature16961

Simpson, G.G., 1984. Tempo and Mode in Evolution. Columbia University Press. https://doi.org/10.7312/simp93040

Slater, G.J., Harmon, L.J., Alfaro, M.E., 2012. Integrating Fossils with Molecular Phylogenies Improves Inference of Trait Evolution. Evolution 66, 3931–3944. https://doi.org/10.1111/j.1558-5646.2012.01723.x

Smart, U., Ingrasci, M.J., Sarker, G.C., Lalremsanga, H., Murphy, R.W., Ota, H., Tu, M.C., Shouche, Y., Orlov, N.L., Smith, E.N., 2021. A comprehensive appraisal of evolutionary diversity in venomous Asian coralsnakes of the genus Sinomicrurus (Serpentes: Elapidae) using Bayesian coalescent inference and supervised machine learning. Journal of Zoological Systematics and Evolutionary Research 59, 2212–2277. https://doi.org/10.1111/jzs.12547

Smith, M.L., Hahn, M.W., 2023. Phylogenetic inference using generative adversarial networks. Bioinformatics 39, btad543. https://doi.org/10.1093/bioinformatics/btad543

Smith, N.D., Turner, A.H., 2005. Morphology's Role in Phylogeny Reconstruction: Perspectives from Paleontology. Systematic Biology 54, 166–173. https://doi.org/10.1080/10635150590906000

Söderkvist, O.J.O., 2016. Swedish Leaf Dataset.

Söderkvist, O.J.O., 2001. Computer vision classifcation of leaves from swedish trees (Master's Thesis). Linkoping University.

Soltis, P.S., Nelson, G., Zare, A., Meineke, E.K., 2020. Plants meet machines: Prospects in machine learning for plant biology. Appl Plant Sci 8, e11371. https://doi.org/10.1002/aps3.11371

Sosiak, C.E., Barden, P., 2021. Multidimensional trait morphology predicts ecology across ant lineages. Functional Ecology 35, 139–152. https://doi.org/10.1111/1365-2435.13697

Spradley, J.P., Glazer, B.J., Kay, R.F., 2019. Mammalian faunas, ecological indices, and machine-learning regression for the purpose of paleoenvironment reconstruction in the Miocene of South America. Palaeogeography, Palaeoclimatology, Palaeoecology 518, 155–171. https://doi.org/10.1016/j.palaeo.2019.01.014

Srivastava, A., Klassen, E., Joshi, S.H., Jermyn, I.H., 2011. Shape Analysis of Elastic Curves in Euclidean Spaces. IEEE Trans. Pattern Anal. Mach. Intell. 33, 1415–1428. https://doi.org/10.1109/TPAMI.2010.184

Stanford University, H., 2022. Artificial Intelligence Index Report 2022.

Sun, C.-Y., Yang, Y.-Q., Guo, H.-X., Wang, P.-S., Tong, X., Liu, Y., Shum, H.-Y., 2023. Semi-supervised 3D shape segmentation with multilevel consistency and part substitution. Comp. Visual Media 9, 229–247. https://doi.org/10.1007/s41095-022-0281-9

Sun, X., Shi, J., Liu, L., Dong, J., Plant, C., Wang, X., Zhou, H., 2018. Transferring deep knowledge for object recognition in Low-quality underwater videos. Neurocomputing 275, 897–908. https://doi.org/10.1016/j.neucom.2017.09.044

Suvorov, A., Hochuli, J., Schrider, D.R., 2020. Accurate Inference of Tree Topologies from Multiple Sequence Alignments Using Deep Learning. Systematic Biology 69, 221–233. https://doi.org/10.1093/sysbio/syz060

Tan, H., Qiu, S., Wang, J., Yu, G., Guo, W., Guo, M., 2022. Weighted deep factorizing heterogeneous molecular network for genome-phenome association prediction. Methods 205, 18–28. https://doi.org/10.1016/j.ymeth.2022.05.008

Tang, X., Yuan, Y., Li, X., Zhang, J., 2021. Maximum Entropy Modeling to Predict the Impact of Climate Change on Pine Wilt Disease in China. Frontiers in Plant Science 12. https://doi.org/0.3389/fpls.2021.652500

Thermo Fisher Scientific, 2021. Avizo.

Toussaint, N., Redhead, Y., Vidal-García, M., Lo Vercio, L., Liu, W., Fisher, E.M.C., Hallgrímsson, B., Tybulewicz, V.L.J., Schnabel, J.A., Green, J.B.A., 2021. A landmark-free morphometrics pipeline for high-resolution phenotyping: application to a mouse model of Down syndrome. Development 148, dev188631. https://doi.org/10.1242/dev.188631

Tseng, Z.J., Garcia-Lara, S., Flynn, J.J., Holmes, E., Rowe, T.B., Dickson, B.V., 2023. A switch in jaw form–function coupling during the evolution of mammals. Phil. Trans. R. Soc. B 378, 20220091. https://doi.org/10.1098/rstb.2022.0091

Tuladhar, A., Schimert, S., Rajashekar, D., Kniep, H.C., Fiehler, J., Forkert, N.D., 2020. Automatic Segmentation of Stroke Lesions in Non-Contrast Computed Tomography Datasets With Convolutional Neural Networks. IEEE Access 8, 94871–94879. https://doi.org/10.1109/ACCESS.2020.2995632

Ubbens, J., Cieslak, M., Prusinkiewicz, P., Parkin, I., Ebersbach, J., Stavness, I., 2020. Latent Space Phenotyping: Automatic Image-Based Phenotyping for Treatment Studies. Plant Phenomics 2020. https://doi.org/10.34133/2020/5801869

Unger, S., Rollins, M., Tietz, A., Dumais, H., 2021. iNaturalist as an engaging tool for identifying organisms in outdoor activities. Journal of Biological Education 55, 537–547. https://doi.org/10.1080/00219266.2020.1739114

Valan, M., Makonyi, K., Maki, A., Vondráček, D., Ronquist, F., 2019. Automated Taxonomic Identification of Insects with Expert-Level Accuracy Using Effective Feature Transfer from Convolutional Networks. Systematic Biology 68, 876–895. https://doi.org/10.1093/sysbio/syz014

Van Der Bijl, W., Zeuss, D., Chazot, N., Tunström, K., Wahlberg, N., Wiklund, C., Fitzpatrick, J.L., Wheat, C.W., 2020. Butterfly dichromatism primarily evolved via Darwin's, not Wallace's, model. Evolution Letters 4, 545–555. https://doi.org/10.1002/evl3.199

van der Walt, S., Schönberger, J.L., Nunez-Iglesias, J., Boulogne, F., Warner, J.D., Yager, N., Gouillart, E., Yu, T., contributors, the scikit-image, 2014. scikit-image: image processing in Python. PeerJ 2, e453. https://doi.org/10.7717/peerj.453

Vasconcelos, T., Boyko, J.D., Beaulieu, J.M., 2023. Linking mode of seed dispersal and climatic niche evolution in flowering plants. Journal of Biogeography 50, 43–56. https://doi.org/10.1111/jbi.14292

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L., Polosukhin, I., 2017. Attention Is All You Need. CoRR abs/1706.03762.

Vedaldi, A., Favaro, P., Grisan, E., 2007. Boosting Invariance and Efficiency in Supervised Learning, in: 2007 IEEE 11th International Conference on Computer Vision. Presented at the 2007 IEEE 11th International Conference on Computer Vision, IEEE, Rio de Janeiro, Brazil, pp. 1–8. https://doi.org/10.1109/ICCV.2007.4408840

Vermillion, W.A., Polly, P.D., Head, J.J., Eronen, J.T., Lawing, A.M., 2018. Ecometrics: A Trait-Based Approach to Paleoclimate and Paleoenvironmental Reconstruction, in: Croft, D.A., Su, D.F., Simpson, S.W. (Eds.), Methods in Paleoecology, Vertebrate Paleobiology and Paleoanthropology. Springer International Publishing, Cham, pp. 373–394. https://doi.org/10.1007/978-3-319-94265-0_17

Viroli, C., McLachlan, G.J., 2019. Deep Gaussian mixture models. Stat Comput 29, 43–51. https://doi.org/10.1007/s11222-017-9793-z

Vurtur Badarinath, P., Chierichetti, M., Davoudi Kakhki, F., 2021. A Machine Learning Approach as a Surrogate for a Finite Element Analysis: Status of Research and Application to One Dimensional Systems. Sensors 21, 1654. https://doi.org/10.3390/s21051654

Vydana, H.K., Karafiat, M., Zmolikova, K., Burget, L., Cernocky, H., 2021. Jointly Trained Transformers Models for Spoken Language Translation, in: ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Presented at the ICASSP 2021 - 2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, Toronto, ON, Canada, pp. 7513–7517. https://doi.org/10.1109/ICASSP39728.2021.9414159

Wäldchen, J., Mäder, P., 2018. Machine learning for image based species identification. Methods in Ecology and Evolution 9, 2216–2225. https://doi.org/10.1111/2041-210X.13075

Walker, B.E., Tucker, A., Nicolson, N., 2022. Harnessing Large-Scale Herbarium Image Datasets Through Representation Learning. Frontiers in Plant Science 12, 1–12. https://doi.org/10.3389/fpls.2021.806407

Wang, L., Shao, J., Fang, F., 2021. Propensity model selection with nonignorable nonresponse and instrument variable. STAT SINICA. https://doi.org/10.5705/ss.202019.0025

Wang, Y., Yao, Q., Kwok, J.T., Ni, L.M., 2021. Generalizing from a Few Examples: A Survey on Few-shot Learning. ACM Comput. Surv. 53, 1–34. https://doi.org/10.1145/3386252

Weaver, W.N., Ng, J., Laport, R.G., 2020. LeafMachine: Using machine learning to automate leaf trait extraction from digitized herbarium specimens. Applications in Plant Sciences 8, e11367. https://doi.org/10.1002/aps3.11367

Weller, H.I., Van Belleghem, S.M., Hiller, A.E., Lord, N.P., 2022. Flexible color segmentation of biological images with the R package recolorize (preprint). Bioinformatics. https://doi.org/10.1101/2022.04.03.486906

Whang, S.E., Roh, Y., Song, H., Lee, J.-G., 2023. Data collection and quality challenges in deep learning: a data-centric AI perspective. The VLDB Journal 32, 791–813. https://doi.org/10.1007/s00778-022-00775-9

What's next for bioimage analysis?, 2023. . Nat Methods 20, 945–946. https://doi.org/10.1038/s41592-023-01950-8

White, A.E., Dikow, R.B., Baugh, M., Jenkins, A., Frandsen, P.B., 2020. Generating segmentation masks of herbarium specimens and a data set for training segmentation models using deep learning. Applications in Plant Sciences 8, 1–8. https://doi.org/10.1002/aps3.11352

White, H.E., Goswami, A., Tucker, A.S., 2021. The Intertwined Evolution and Development of Sutures and Cranial Morphology. Front Cell Dev Biol 9, 653579. https://doi.org/10.3389/fcell.2021.653579

Wiens, J.J., 2006. Missing data and the design of phylogenetic analyses. Journal of Biomedical Informatics 39, 34–42. https://doi.org/10.1016/j.jbi.2005.04.001

Wiens, J.J., 2001. Character Analysis in Morphological Phylogenetics: Problems and Solutions. Systematic Biology 50, 689–699. https://doi.org/10.1080/106351501753328811

Wilf, P., Wing, S.L., Meyer, H.W., Rose, J.A., Saha, R., Serre, T., Cúneo, N.R., Donovan, M.P., Erwin, D.M., Gandolfo, M.A., González-Akre, E., Herrera, F., Hu, S., Iglesias, A., Johnson, K.R., Karim, T.S., Zou, X., 2021. An image dataset of cleared, x-rayed, and fossil leaves vetted to plant family for human and machine learning. PhytoKeys 187, 93–128. https://doi.org/10.3897/PHYTOKEYS.187.72350

Willers, C., Bauman, G., Andermatt, S., Santini, F., Sandkühler, R., Ramsey, K.A., Cattin, P.C., Bieri, O., Pusterla, O., Latzin, P., 2021. The impact of segmentation on whole-lung functional MRI quantification: Repeatability and reproducibility from multiple human observers and an artificial neural network. Magnetic Resonance in Med 85, 1079–1092. https://doi.org/10.1002/mrm.28476

Wilson, R.J., De Siqueira, A.F., Brooks, S.J., Price, B.W., Simon, L.M., Van Der Walt, S.J., Fenberg, P.B., 2023. Applying computer vision to digitised natural history collections for climate change research: Temperature-size responses in British butterflies. Methods Ecol Evol 14, 372–384. https://doi.org/10.1111/2041-210X.13844

Withers, P.J., Bouman, C., Carmignato, S., Cnudde, V., Grimaldi, D., Hagen, C.K., Maire, E., Manley, M., Du Plessis, A., Stock, S.R., 2021. X-ray computed tomography. Nat Rev Methods Primers 1, 18. https://doi.org/10.1038/s43586-021-00015-4

Wolfram Research, Inc., 2024. Mathematica.

Wu, Di, Wu, Dan, Feng, H., Duan, L., Dai, G., Liu, X., Wang, K., Yang, P., Chen, G., Gay, A.P., Doonan, J.H., Niu, Z., Xiong, L., Yang, W., 2021. A deep learning-integrated micro-CT image analysis pipeline for quantifying rice lodging resistance-related traits. Plant Communications 2, 100165. https://doi.org/10.1016/j.xplc.2021.100165

Wu, Y., Kirillov, A., 2019. Detectron2.

Wu, Z., Xiong, Y., Yu, S., Lin, D., 2018. Unsupervised Feature Learning via Non-Parametric Instance-level Discrimination.

Yang, B., Zhang, Z., Yang, C.-Q., Wang, Y., Orr, M.C., Wang, H., Zhang, A.-B., 2022. Identification of Species by Combining Molecular and Morphological Data Using Convolutional Neural Networks. Systematic Biology 71, 690–705. https://doi.org/10.1093/sysbio/syab076

Yang, Z., 2015. The BPP program for species tree estimation and species delimitation. Current

Zoology 61, 854–865. https://doi.org/10.1093/czoolo/61.5.854

Ye, S., Lauer, J., Zhou, M., Mathis, A., Mathis, M.W., 2023. AmadeusGPT: a natural language interface for interactive animal behavioral analysis. https://doi.org/10.48550/arXiv.2307.04858

Young, R., Maga, A.M., 2015. Performance of single and multi-atlas based automated landmarking methods compared to expert annotations in volumetric microCT datasets of mouse mandibles. Front Zool 12, 33. https://doi.org/10.1186/s12983-015-0127-8

Yu, C., Qin, F., Li, Y., Qin, Z., Norell, M., 2022. CT Segmentation of Dinosaur Fossils by Deep Learning. Front. Earth Sci. 9, 805271. https://doi.org/10.3389/feart.2021.805271

Yu, C., Qin, F., Watanabe, A., Yao, W., Li, Y., Qin, Z., Liu, Y., Wang, H., Jiangzuo, Q., Hsiang, A.Y., Ma, C., Rayfield, E., Benton, M.J., Xu, X., 2023. AI in paleontology. https://doi.org/10.1101/2023.08.07.552217

Yu, L., Shi, J., Huang, C., Duan, L., Wu, D., Fu, D., Wu, C., Xiong, L., Yang, W., Liu, Q., 2021. An integrated rice panicle phenotyping method based on X-ray and RGB scanning and deep learning. The Crop Journal 9, 42–56. https://doi.org/10.1016/j.cj.2020.06.009

Zaharias, P., Grosshauser, M., Warnow, T., 2022. Re-evaluating Deep Neural Networks for Phylogeny Estimation: The Issue of Taxon Sampling. Journal of Computational Biology 29, 74–89. https://doi.org/10.1089/cmb.2021.0383

Zarkogiannis, S.D., Antonarakou, A., Fernandez, V., Mortyn, P.G., Kontakiotis, G., Drinia, H., Greaves, M., 2020a. Evidence of Stable Foraminifera Biomineralization during the Last Two Climate Cycles in the Tropical Atlantic Ocean. JMSE 8, 737. https://doi.org/10.3390/jmse8100737

Zarkogiannis, S.D., Kontakiotis, G., Gkaniatsa, G., Kuppili, V.S.C., Marathe, S., Wanelik, K., Lianou, V., Besiou, E., Makri, P., Antonarakou, A., 2020b. An Improved Cleaning Protocol for Foraminiferal Calcite from Unconsolidated Core Sediments: HyPerCal—A New Practice for Micropaleontological and Paleoclimatic Proxies. JMSE 8, 998. https://doi.org/10.3390/jmse8120998

Zelditch, M.L., Goswami, A., 2021. What does modularity mean? Evolution and Development 23, 377–403. https://doi.org/10.1111/ede.12390

Zeng, A., Yan, L., Huang, Y., Ren, E., Liu, T., Zhang, H., 2021. Intelligent Detection of Small Faults Using a Support Vector Machine. Energies 14, 6242. https://doi.org/10.3390/en14196242

Zhang, D., Maslej, N., Brynjolfsson, E., Etchemendy, J., Lyons, T., Manyika, J., Ngo, H., Niebles, J.C., Sellitto, M., Sakhaee, E., Shoham, Y., Clark, J., Perrault, R., 2022. The AI Index 2022 Annual Report.

Zhang, H., Starke, S., Komura, T., Saito, J., 2018. Mode-adaptive neural networks for quadruped motion control. ACM Trans. Graph. 37, 145:1-145:11. https://doi.org/10.1145/3197517.3201366

Zhang, L., Wang, Y., Ruhl, M., Xu, Y., Zhu, Y., An, P., Chen, H., Yan, D., 2023. Machine-learning-based morphological analyses of leaf epidermal cells in modern and fossil ginkgo and their implications for palaeoclimate studies. Palaeontology 66, e12684. https://doi.org/10.1111/pala.12684

Zhao, M., Liu, Q., Jha, A., Deng, R., Yao, T., Mahadevan-Jansen, A., Tyska, M.J., Millis, B.A., Huo, Y., 2021. VoxelEmbed: 3D Instance Segmentation and Tracking with Voxel Embedding based Deep Learning, in: Lian, C., Cao, X., Rekik, I., Xu, X., Yan, P. (Eds.), Machine Learning in Medical Imaging, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 437–446. https://doi.org/10.1007/978-3-030-87589-3_45

Zhu, S., Brazil, G., Liu, X., 2020. The Edge of Depth: Explicit Constraints Between Segmentation and Depth, in: 2020 IEEE/CVF Conference on Computer Vision and

Pattern Recognition (CVPR). Presented at the 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, Seattle, WA, USA, pp. 13113–13122. https://doi.org/10.1109/CVPR42600.2020.01313

Zhu, X., Goldberg, A.B., 2022. Introduction to semi-supervised learning. Springer Nature.

Zou, Z., Zhang, H., Guan, Y., Zhang, J., 2020. Deep Residual Neural Networks Resolve Quartet Molecular Phylogenies. Molecular Biology and Evolution 37, 1495–1507. https://doi.org/10.1093/molbev/msz307