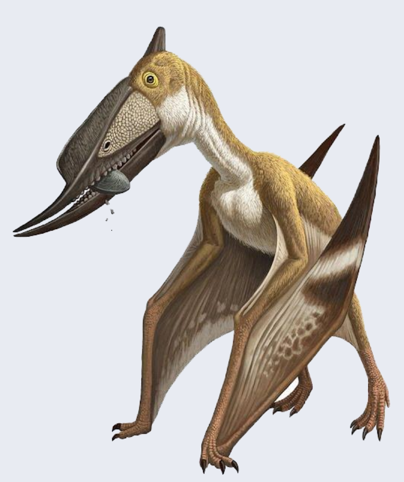
Classifying Dinosaur Tracks Using A Convolutional Neural Net

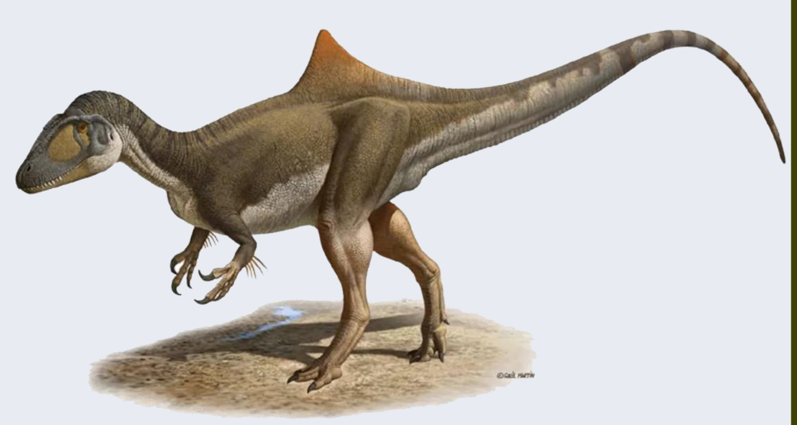
By Adam Maciaszek

Introduction

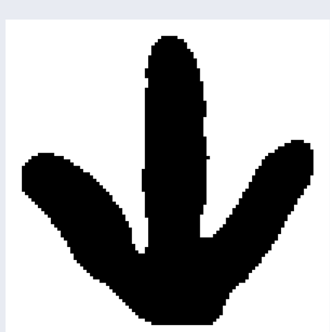
Amherst, Massachusetts has a significant connection to ichnology, the study of trace fossils largely due to the work of Edward Hitchcock, a 19th-century geologist and paleontologist who lived and worked in Amherst. Hitchcock developed a classification system for footprints based on their shape and other characteristics. The fossilized impressions left in the mud tell important information about the species that walked there millions of years ago such as their gate and patterns of locomotion and their place in the evolution tree. Unfortunately, it can be very difficult to identify the animals that left the tracks there between erosion and impartial footprint impressions left. There are two major groups of dinosaurs Saurischians and Ornithischians which would evolve into birds and the other to eventually become extinct.

Distinguishing between tridactyl three-toed dinosaur tracks of the herbivorous ornithischians and the predominantly carnivorous theropods is a complex and long-standing problem. Broadly, ornithischian tracks are expected to be wider and more symmetric than theropod tracks . Identifying which tracks belong to prey and predator is very useful since there are many tracks left behind by therapods in pursuit of prey. These three toed ornithischians tracks are one of the most difficult to distinguish between even for trained paleontologists trained in the field. The aim of this project is to use an existing online database of dinosaur track traces to identify tracks into ornithischians and therapods with three toes.

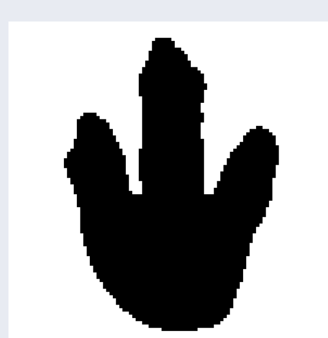




Ornithischian

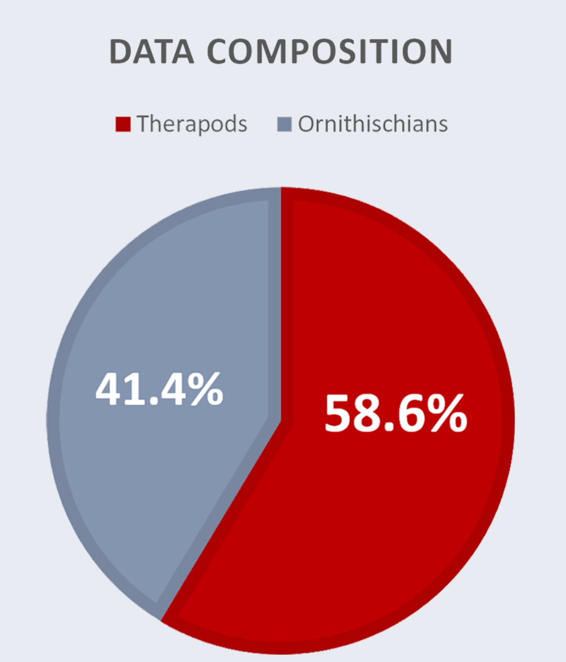


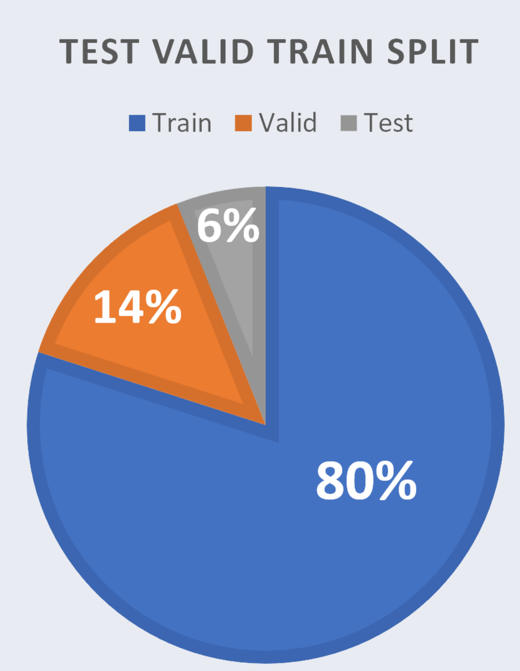
Therapod



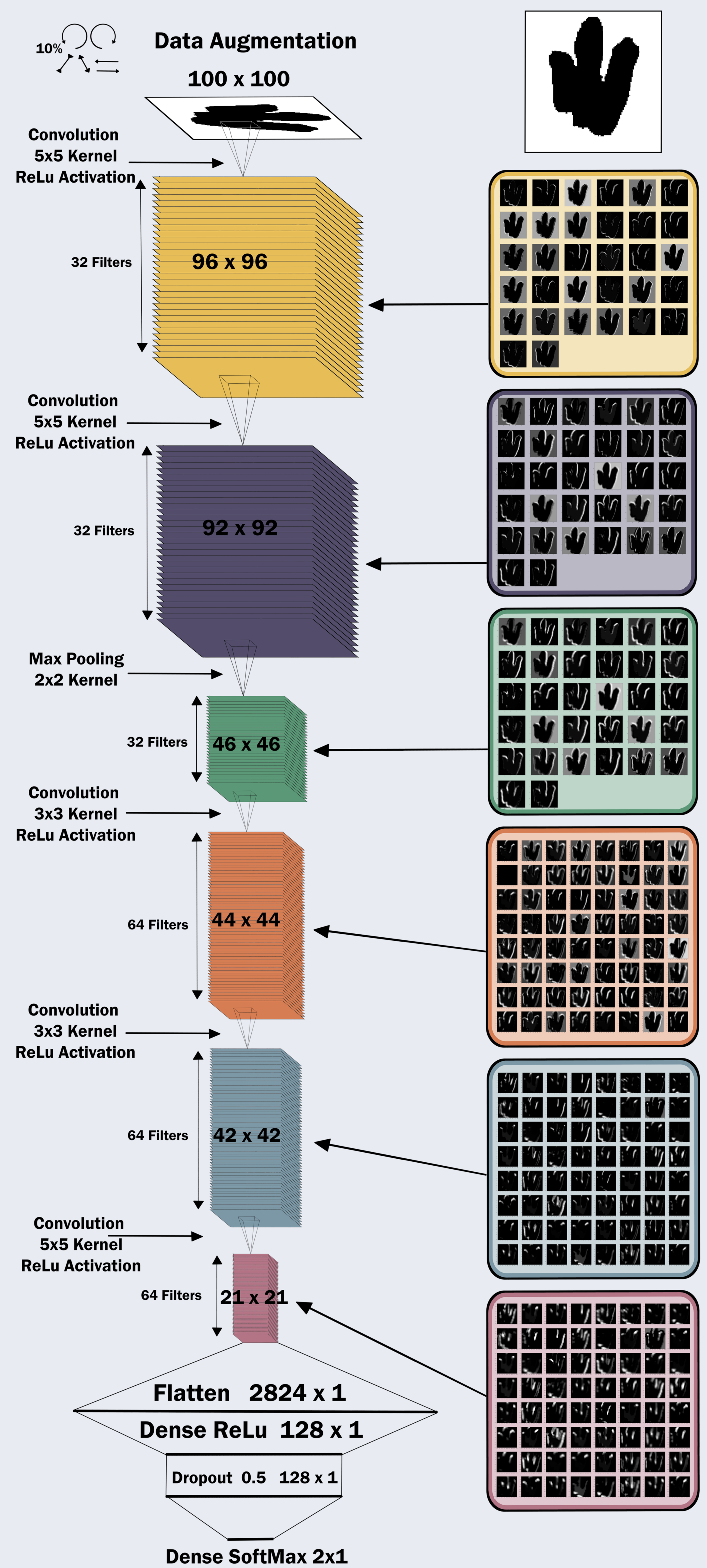
The Australian Dataset

The dataset consists of monochromatic traces with a resolution of 100 by 100 pixels, limited to black and white pixel values to minimize noise. The data is divided into 58.6% therapods and 41.4% Ornithischians. The trace outlines were created by Jens Lallensack using photogrammetry data, which includes depth information. These traces were obtained from the Lark Quarry trackway in Australia, also known as the Dinosaur Stampede National Monument





Model Design



Data augmentation

- Introduce variations in the training data.
- Transformations applied: rotation (±10 degrees), shifting (±10 pixels), and scaling (±10%)
- Data augmentation enhances the model's robustness and its ability to generalize to unseen data.

Convolutional Layers Conv2D

- Apply learnable filters to input image
- Generate feature maps capturing local features
- Determine receptive field size
- Capture intricate details in the image
- Two initial Conv2D layers ~ 64 filters with a kernel size of 5x5
- Next two Conv2D layers ~ 64 filters with a kernel size of 3x3

ReLU Activation function (Rectified Linear Unit)

- Applied in Conv2D and Dense layers
- Adds non-linearity to the network
- Neuron output is input value if positive, zero if negative
- High computational efficiency and avoids vanishing gradient problem
- Enables learning and classification of complex features
- Enhances accuracy of feature extraction and classification

Pooling layer (MaxPool2D)

- Reduces spatial dimensions of Conv2D feature maps enhancing efficiency and performance
- Downsampling reduces computational complexity and improves generalization
- MaxPool2D selects maximum value within neighboring pixels Pooling window size determines extent of downsampling
- MaxPool2D preserves salient features while reducing dimensions
- Mitigates overfitting in the model

Flatten layer

- Converts multidimensional feature maps to a one-dimensional vector
- Connects CNN with fully connected layers for classification
- Facilitates extraction of high-level information from the image
- Enhances classification accuracy

Dropout layer

- Regularization technique that encourages distributed feature learning
- Deactivates nodes randomly during training
- Dropout rate hyperparameter determines node deactivation proportion
- Improves generalization performance and mitigates overfitting

Dense layers

- Two used for feature extraction from flattened feature maps.
- First Dense layer with 128 neurons to capture intricate patterns and representations.
- Second Dense layer with 2 neurons for binary classification.
- Final layer utilizes SoftMax activation. Outputs a probability distribution over the classes.
- Enables intuitive interpretation and decision-making based on highest probability.

Learning Rate Annealer

- Fine-tuning the learning process enhances model accuracy and overall performance.
- Gradually decreasing the learning rate during training improves convergence.

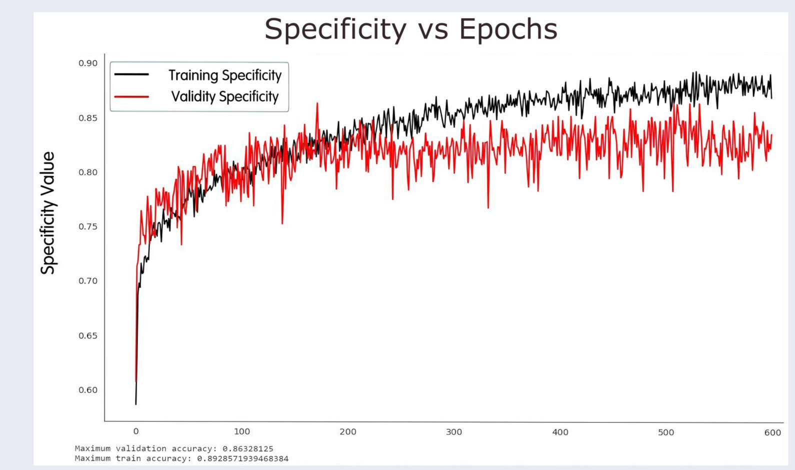
Metric: Specificity

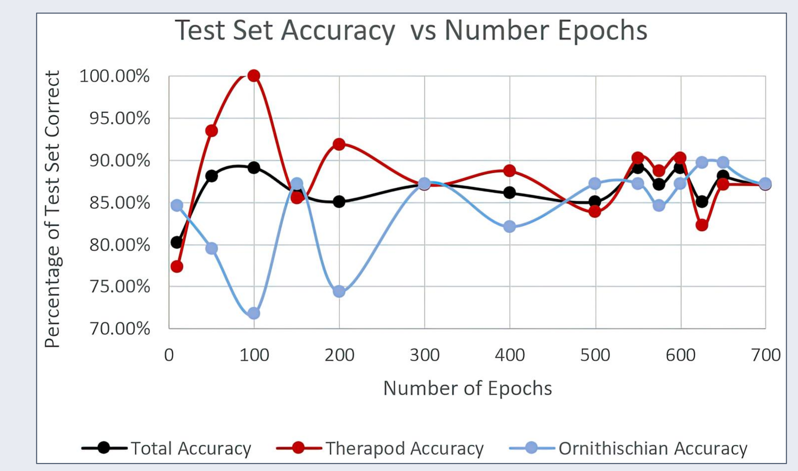
- Measures the model's ability to identify true negatives in binary classification
- Reduces false positive predictions, and improves accuracy of identifying true negatives
- Calculated as the ratio of true negatives to the sum of true negatives and false positives
- Valuable when the negative class is important or false positives are undesirable
- Important considering the negative class comprises only 41% of the data
- Alternative metrics like f1 or accuracy yield lower accuracy for the negative class.

Results

- Deciding how many Epochs was a difficult decision since the point where training accuracy and validation converge based on 100 to 150 epochs
- At 100 epochs there is the highest over all accuracy which is a very misleading 90% accuracy since therapods prediction is 100% for the test set but only 56% correct for ornithischians the best-balanced results found at 600 Epochs
- Higher overall accuracy was achieved with the scoring metric for accuracy rather than sensitivity resulting in 92% overall accuracy in the test set versus 89% total accuracy when sensitivity id the scoring metric. This results in a very biased model when accuracy is the scoring metric with 95% accuracy for Therapods but only 61% for ornithischians under best conditions

Metric Best Case	Overall Accuracy	Therapod Accuracy	Ornithischian Accuracy
Accuracy	91%	100%	79%
Specificity	89%	90%	87%





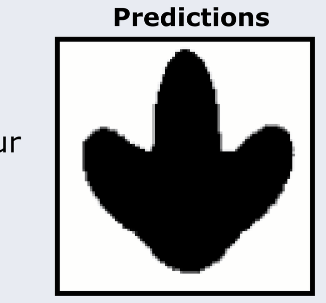
Results Compared to Paleontologists

Identifier	Correct	Ambiguous	Incorrect
Expert 1	67%	3%	31%
Expert 2	58%	25%	17%
Expert 3	58%	25%	17%
Expert 4	42%	44%	14%
Expert 5	58%	22%	19%
Previous Model No Ambiguous Cases	86%	0%	14%
This Model No Ambigous Cases	89%	0%	11%
Previous Model 0.4 - 0.6 treated as ambiguous	67%	22%	11%
This Model 0.4 - 0.6 treated as ambiguous	89%	1%	10%

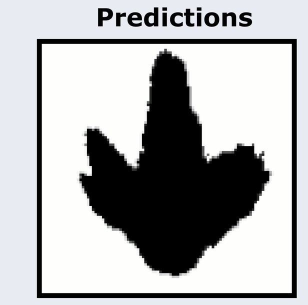
- On the table to the left you can see how both machine learning models vastly outperformed even the best of the experts sampled who at best got 67% correct versus 86% of previous research and the 89% of this model
- Aside from the 3% increase in accuracy there is a much larger increase in confidence in predictions. While the previous research found 22% images to be too ambiguous with a probability between 0.4 to 0.6 while this model found only 1% of the images to be too ambiguous.

Median Of All Pixel Values

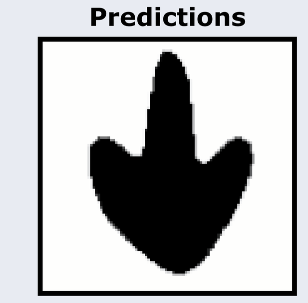
True Ornithischian Predictions



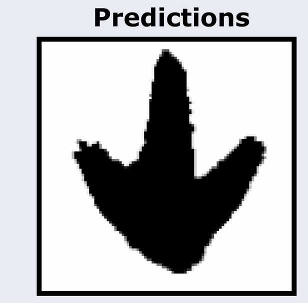
False Ornithischian Predictions



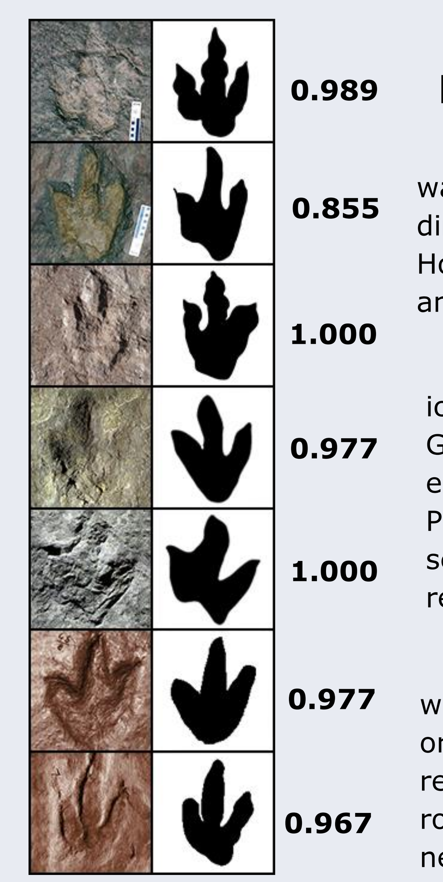
True Therapod Predictions



False Therapod Predictions



- The model categorized all images in the set sorted them into four categories: correct and incorrect predictions for theropods, and correct and incorrect predictions for ornithischians. Median pixel values were calculated for each position.
- For the correct predictions, ornithischians exhibit wider and symmetrical tracks, while theropods have narrower tracks with a significantly longer central digit.
- The misidentified therapods, on average, displayed wider tracks, while the misidentified ornithischians all had considerably longer central digits.



Massachusetts Self Collected Data

To assess the reliability and generalization of the model it was tested on local tracks. The local tracks I took photos at dinosaur footprint site on route 5 in Holyoke and at Wisteria Hurst Holyoke and I also used lithographs from the Beneski Museum and were local tracks sourced by Edward Hitchcock.

These local fossils were all Eubrontes Giganteus an ichnospecies designated by Hitchcock. We know Eubrontes Giganteus is late Triassic origin and is a theropod but not the exact species. The photos were cropped and centered using Photoshop, and an outline of the track was drawn in black on a separate layer and filled. The image was down sampled to a resolution of 100x100 pixels for model evaluation.

All 13 tracks were accurately identified as theropod traces with a minimum probability of 0.85. However, due to the lack of ornithischian data, the model's performance on external datasets remains inconclusive. Nevertheless, the model demonstrated robustness in theropod identification, without reporting any false negatives.