A Convolutional Neural Network Classification of Species within the Genus *Carcharhinus*

Mikey Sison

13 May 2023

Abstract

Over the past few years, convolutional neural networks (CNN) and deep learning networks have dominated the forefront of image classification and recognition, significantly out-classing artificial neural networks. Over the same period, morphometric analyses have yet to adopt these new technologies. We aim to utilize these new techniques to classify the lateral teeth of nine modern shark species of the genus *Carcharhinus*. Prior research has utilized elliptical Fourier analysis coupled with discriminant analysis and artificial neural networks to identify biological shapes. These analyses focused on numerical morphological characteristics. Our analysis begins directly from the images, superseding the need to manually collect the characteristics. We push the bounds of supervision by applying newer techniques to unearth features and relationships previously undetected among teeth. A two-step baseline method will be used to incrementally test the classification accuracy of the models. We create three primary models: the first being a SVC (Support Vector Classifier) using ORB (Rublee et al., 2011) feature extraction, the second being a pretrained Convolutional Neural Network– VGG (Very Deep Convolutional Network), and the third being a patch based VGG implementation. It is assumed that the classification error will decrease with increasing model complexity along these bounds, with the CNN being more complex than the SVC. To verify that this assumption is correct, the previous classification accuracy was used as a baseline for the more complex model to beat. Previous research has boasted a classification accuracy of 40-55% over the nine species classes. With the three methods employed in the paper, we achieved classification accuracies of 52%, 61%, and 83.5%.

Introduction

Sharks are some of the most well known and abundant marine predators of the modern oceans. These predators often dominate the highest trophic levels, with some being the top predator by biomass in certain oceans (Sandin et al., 2022). Sharks are divided into eight orders, the most prolific of which is the Carcharhiniformes. Carcharhiniformes species number in the hundreds, containing the majority of the shark species that rule the oceans today. The Carcharhiniformes are broken down into eight families, which include the family Carcharhinidae. This family houses the genus *Carcharhinus*, which include about fifty species, found in all ocean basins (Dosay-Akbulut, 2008; Naylor, 1992). Many of the species are quite well known by the public, including the bull shark (*Carcharhinus leucas*) or the blacktip reef shark (*Carcharhinus melanopterus*). *Carcharhinus* has a long history of being a top predator. This history can be best examined by exploring the fossil record.

Although there is a well-documented fossil record, the evolution of *Carcharhinus* is not well understood (Purdy et al., 2001; Smith, 2015). The full phylogenetic tree of the genus *Carcharhinus* is highly debated with organization and membership of the genus disputed with many species of *Carcharhinids* being proposed over the past few decades (Castro, 2011; Dosay-Akbulut, 2008; Harry et al. 2012). The fossil record of *Carcharhinus* and most shark species are quantified almost entirely by their teeth. While other animals like dinosaurs leave behind fossilized portions of their bony skeleton, shark skeletons consist almost entirely of calcified cartilage. Long et al. (2015) found that sharks evolved from more bony ancestors, implying that the evolution of a lighter cartilaginous endoskeleton improved their predatory abilities. The lighter and more agile body frame is a key component of what drove their domination and proliferation throughout all modern and historic marine ecosystems and the fossil record. Unfortunately, this cartilaginous structure is less durable to weathering and rapidly decomposes after the death of the shark . Due to this rapid decomposition, the primary fossil found that can distinguish shark taxa is teeth. The rapid production of shark teeth throughout a lifetime, ranging in the thousands per shark, and the sturdier enamel of shark teeth contribute further to this fact (Seidel et al., 2017; Kemp and Westrin, 1979). Because we only have fossil teeth to categorize the evolution of *Carcharhinus*, the teeth’s morphometric similarities and dental positional differences make it difficult to distinguish between species in the genus. Highlighted in figure 1, we can see the interspecies similarities between three species within the genus. Creating a framework to effectively and accurately distinguish the species of *Carcharhinus* given fossil teeth will vastly improve our understanding of the evolution and paleobiology of *Carcharhinus* species.



Figure 1: A sample of three species of *Carcharhinus*, highlighting their visual similarity. From left to right, Bull shark (*Carcharhinus leucas*), Dusky shark (*Carcharhinus obscurus*), Oceanic Whitetip (*Carcharhinus longimanus*).

The classification of shark teeth of the genus *Carcharhinus* is a notoriously difficult problem with recent methods only resulting in 40-55% accuracy between the nine species in question (Smith, 2015) and similar methods achieving ~90% accuracy but only between four species (Soda, 2013). *Carcharhinids'* visually similar teeth structures make it extremely difficult to differentiate between species using visual cues or heuristics. Current methodologies primarily employ landmark-based geometric morphometrics and outline-based morphometrics, which apply elliptic Fourier analysis (EFA) and discriminant analysis to classify teeth. These methods, while effective, could be failing to capture features intrinsic to the tooth image. We aim to test this with a new approach to classifying shark teeth based on newer and more robust learned classification regimes.

Current classification regimes focus on landmark-based geometric morphometrics and outline-based morphometrics (Naylor & Marcus, 1994; Smith, 2015; Soda, 2013). In essence, these analyses rely upon manual measurements making them extremely time and resource intensive. Although an analysis done by Soda (2013) achieved a +90% accuracy using an artificial neural network, our analysis differs in a few key ways. First, their analysis and those done in Smith (2015) used geometric and morphometric markers instead of images. Soda (2013) only classified *C. acronotus, C. leucas, C. limbatus,* and *C. plumbeus* instead of a larger sample of species which we will use (see appendix 1). Using the image data from Smith (2015), we selected nine species from the genus *Carcharhinus*, outlined in appendix 1, to perform our analysis. Using images, we will also be able to discriminate on factors other than the simple shape/outline of the tooth. We will be able to use features such as the dimensions of the root or the neck that might not be analyzed in landmark-based analyses. These features can be difficult to capture using EFA as it only highlights the outline of the tooth.

Since morphologic analyses are used in a wide variety of biological and geological realms (Hirshfield & Reegs, 1978; Bondesan et al., 1992; Smith, 1998; Bookstein et al., 1999), this project has the potential to shift the analysis regime away from a somewhat tedious analysis method of geometric morphometrics to a potentially more accurate and time-efficient method. We hypothesize that our classification regime will classify these nine species of *Carcharhinus* with significantly higher classification accuracy than Smith (2015). To do so, we will first create a baseline Support Vector Machine (SVM) to judge the effectiveness of our Convolutional Neural Network (CNN) and patch-CNN to a simpler classification regime. We expect that the pretrained CNN will be able to better classify the sample teeth than the SVM and Smith (2015), and the patch-based CNN is expected to work best.

Methods

*Data Sourcing*

The images from Smith (2015) were the primary source of data among other images from Dr. Jeffery Agnew’s personal collection. Each image from Dr. Agnew’s collection was taken in accordance with the methods used in Smith (2015). Smith (2015) found that the posterior and anterior teeth were more likely to be misclassified and confounded the overall species classification. Soda (2013) also found that certain inaccuracies arise when including teeth from certain regions of the jaw and sought to account for that. Due to this, we opted to remove the posterior and anterior images from the dataset choosing to focus on the lateral teeth, which had higher classification accuracy. This meant we removed images of the first two tooth positions as well as any above position ten. This was done after consulting classification accuracies by tooth position in Smith (2015). For each class, there are between 25 and 50 images. These images were split into training, test, and validation sets of 55%, 20%, and 25%, respectively, additionally stratifying the splits in order to maintain an equal ratio of classes.

*“Bag of Visual Words”: an SVC Approach*

Using the ~55% classification accuracy from Smith (2015) as a baseline, we decided to implement a novel approach to incrementally improving upon the ~55% baseline. First, to test the accuracy of a simpler model, we are building a model utilizing the “bag of visual words” principle. The “bag of words” principle is adopted from the neighboring field of Natural Language Processing which aims to automatically discern meaning and sentiment behind typed words and sentences. It does so by counting the number of appearances of various words then running a machine learning model to cluster the words into groups (Joachims, 1998). The “bag of visual words” principle works similarly. It works by using SIFT to extract “features” or patches of the image which represent some interesting characteristic of the image such as an edge or a corner, then using the features to represent “words” (Lowe, 1999; Zhang & Zhou, 2010). We then run a K-means clustering algorithm to cluster similar features together. The idea behind this is that we can train a model to become proficient at being able to recognize, for example, specifically, the nutrient groove of the *C. leucas*. Then, if we are able to build up a “bag” of features that represent the features of *C. leucas*, we will be able to accurately classify teeth belonging to *C. leucas*.

In order to determine which features in each bag should belong to which class, we need to apply a support vector machine/classifier. Support Vector Classifiers are machine learning models which attempt to fit a hyperplane to separate classes. The hyperplane has a set of “support vectors” or margins which are optimized to separate the classes with the minimal amount of error / misclassification. A “soft” margin SVC allows for a set threshold of error in order to increase the separability of the classes. With real/non-ideal data, “hard” margins, which disallow error in the creation of the hyperplane, thus assuming linear separability, is highly unrealistic and unlikely to converge (Awad & Khanna, 2015). Thus, we will be using a soft margin SVC to classify and predict on the test set. This method is now slightly outdated and superseded by neural networks, which is why it is being used as a baseline. This method will hereby be referred to as “SVC method” or “Bag of words method”. This method does carry the advantage of being mathematically better suited for smaller datasets vs neural networks. For a deep CNN, with few images in the training and test sets, there is a possibility of overfitting through the convolutions (Awad & Khanna, 2015).

*Data Preprocessing for the CNN Approach*

Prior to running the images through the CNN, the images were preprocessed and resized to ensure uniformity across images. Due to occasional irregular tooth placement within the 5184x3456 resolution image, it was insufficient to simply resize and center crop to the required 224x224. To solve this, we created a program to determine the area the tooth lies and draw a bounding box that captures the tooth in a square crop. This allows us to encapsulate the tooth without stretching or augmenting its dimensions by simply resizing from a rectangle to a square.



Figure 3: (Right) An example of a crop with the solved bounding box.

(Left) An example of center crop problems.

*Choice of CNN*

We originally opted to create our own CNN, but after experimentation, we felt that the dataset proved too small to train our own CNN without overfitting the data. Thus, we opted to use a pretrained CNN named VGG after the Visual Geometry Group (Simonyan & Zisserman, 2014). Pretrained models are advantageous to untrained models in situations with minimal training data. Pretrained models are trained to effectively extract important features from images on massive datasets, often training over millions of images (Iorga & Neagoe, 2019).

*VGG Architecture*

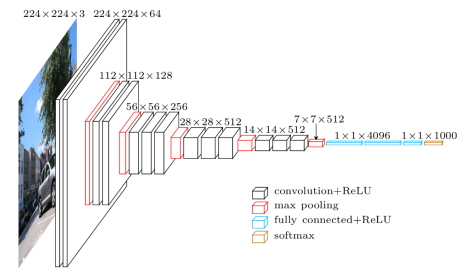
**

Figure 4: Architecture of VGG

The convolutional layers of the pretrained VGG network are designed to extract features from the input image. Each layer uses a set of filters (kernels), learned from the pretrain, to convolve over the input image and produce a set of feature maps. The kernel size of the filters and the stride with which they are applied determines the size of the feature maps. The operation of convolution combines the input of the kernel to produce a new pixel value. Each kernel can be thought of as an n x n matrix that “runs” over the image performing some function (average, gaussian, sharpen, blur) over the pixels. The repeated convolution allows the CNN to start to single out important features of the image.

The pooling layers downsample the feature maps produced by the convolutional layers. By reducing the size of the data, we make the network more efficient. The pooling layers also help to make the network invariant to small changes in the position of objects in the input image. This is extremely important to our dataset. This is the primary way of ensuring stability and consistency between images.

The fully-connected layers predict based on the features extracted by the convolutional layers. These layers take the output of the pooling layers and use it to make a prediction about the input image. Depending on the version in use, the VGG network typically includes one or more fully-connected layers. Next, to determine the prediction accuracies between classes, we will visualize a confusion matrix. This will allow us to visualize which species have lower classification accuracy. Finally, we connect the results to Grad-CAM, which allows us to visualize the feature importances inferred by the CNN (Selvaraju, 2017).

*Patch-based VGG Implementation* From the recent developments in high-resolution medical image classification algorithms have emerged a new regime to manage extremely high resolution images, patch-based CNNs (Hou et. al., 2016; Cruz-Roa et. al., 2014; Mousavi et. al., 2016; Xu et. al., 2015). Since our input images are originally 5184x3456 pixels, we tested the patch-based method. To do so, we first cut the images into patches of 448x448px, double the input size of the CNN, with a stride, or distance between each patch, of 224px. The all black images (pixel values = 0) were subsequently deleted as well. In order to effectively classify each image, a secondary ensemble voting classifier was built on top of the VGG, where each patch from a single image “voted” on the majority species class. Each patch was classified a species and these classifications. The simply majority species then became the overall classified species for a single full input image. These results were then used to train the final VGG CNN.

Results

*SVC Results*

The SVC produced results beyond our expectations with testing accuracies of 61% to 65%. With the best hyper parameters being: regularization term (controls the size of the margin): 1.0, degree: 2, and number of clusters: 9, we achieved training accuracy of ~0.72 and testing accuracy of ~0.61. Already, this implies that the more basic of the two methods surpass the classification rate of methods lined out in Smith, 2015. While SVC is very easy to use, the downside of our current implementation limits our ability to visualize the classification rates of individual classes. For example, it is not currently possible to view how well the SVC classifies the Blacknose shark (*Carcharhinus acronotus*) vs the Oceanic Whitetip (*Carcharhinus longimanus*). This feature is very important in being able to understand which species are commonly misclassified as another. This is paramount to determining the threshold for new species. As discussed, there is much debate about the inclusion of certain species in the genus *Carcharhinus.* Being able to highlight classification accuracy on a species by species scale is key to being able to distinguish between species using this method.

*CNN Results*

Overall, the final CNN achieved an extremely high classification rate of 85.5% testing accuracy. The CNN went through various iterations and compilations to achieve the high classification accuracy. The addition that improved the CNN the most was the usage of the bounding box crop, where the image was cropped according to the limits of the actual tooth itself. This was a great improvement over the previous iteration of the center crop. As highlighted in the confusion matrix of figure 3 and 4, across nearly every class (0-8), the number of correctly classified images increased. For reference, the class labels 0-9 are described in appendix 1. From figure 4, we see that Java shark (*Carcharhinus amboinensis*) and Oceanic Whitetip (*Carcharhinus longimanus*) are frequently misclassified as the Bull shark (*Carcharhinus leucas*). After visual analysis, one can understand as to why this might be the case. The morphology of the teeth are very similar in shape. Naturally, these species would be more difficult to classify, but we believe that this type of misclassification is more a lack of samples than it is a systemic issue.

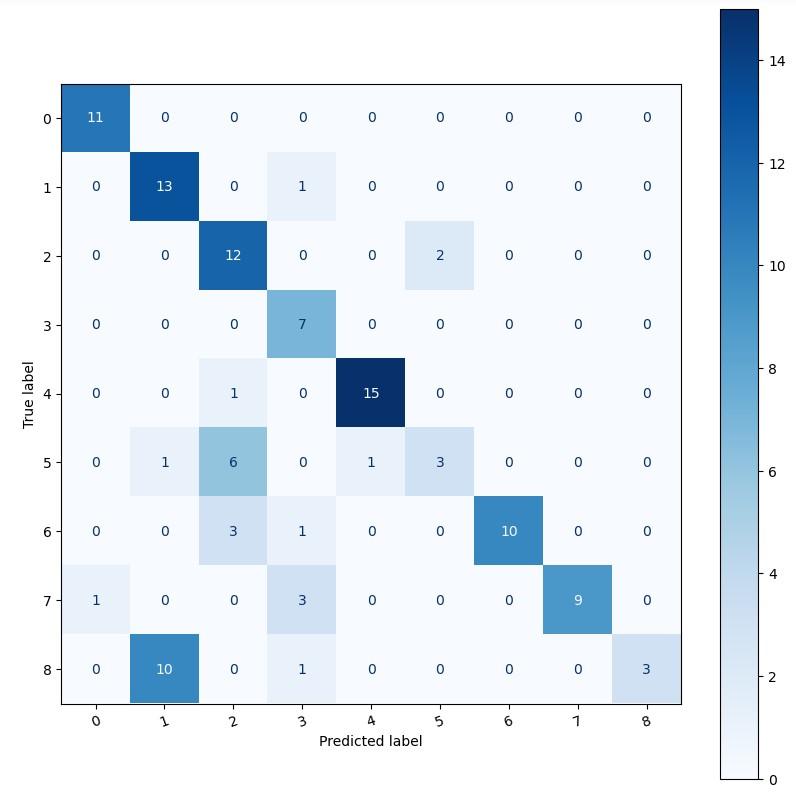


Figure 3: Confusion Matrix of the CNN results using center crop.

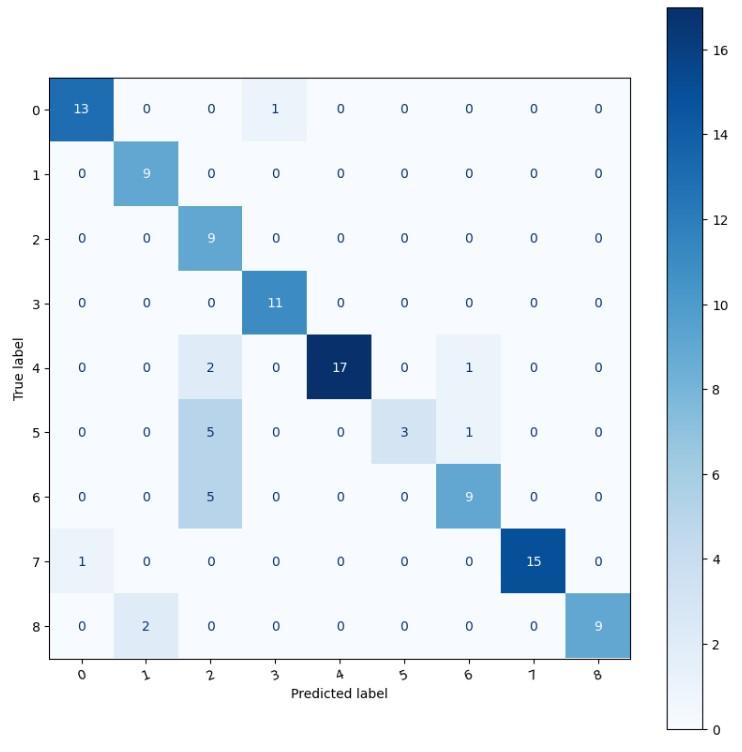


Figure 4: Confusion Matrix of the CNN results using a bounding box.

*Patch-based CNN*

The patch-based CNN suffered a surprisingly lower classification accuracy than expected at a meager 52%. Though Hou 2016 had much better success in a medical field on gigapixel (very large) images, our analysis was less successful. Moreover, the use of the patch-based system disallowed the use of Grad-CAM feature mapping utilized to analyze the most important locations in the CNN implementation. Additionally, the patch system was significantly bulkier and required many more resources to train than the simple VGG. This can be attributed to the significantly higher number of images generated by cutting up the larger image instead of the downsampling used for the CNN. As a result, the lower classification accuracy and higher resource usage was disappointing.

Discussion

*Feature Mapping*

From the CNN, we used a gradient feature map (figure 5) to determine the locations at which the CNN is primarily focusing to distinguish between different classes (Selvaraju, 2017). Interestingly, many of the locations used in the landmark analysis of Naylor & Marcus (1994) are also the locations the CNN uses to classify. For example, Dusky (4) feature map clearly highlights the location “c” from figure 6 (right) and, to a lesser degree, location “e”. Across all feature maps, we see evidence reaffirming the use of these landmarks as differentiating characteristics. Additionally, there are landmarks used by the CNN to classify that are not present on figure 6 that may shed some light on further research on the subject. From the maps, we can conclude the features from each species that are most representative of the species tooth.

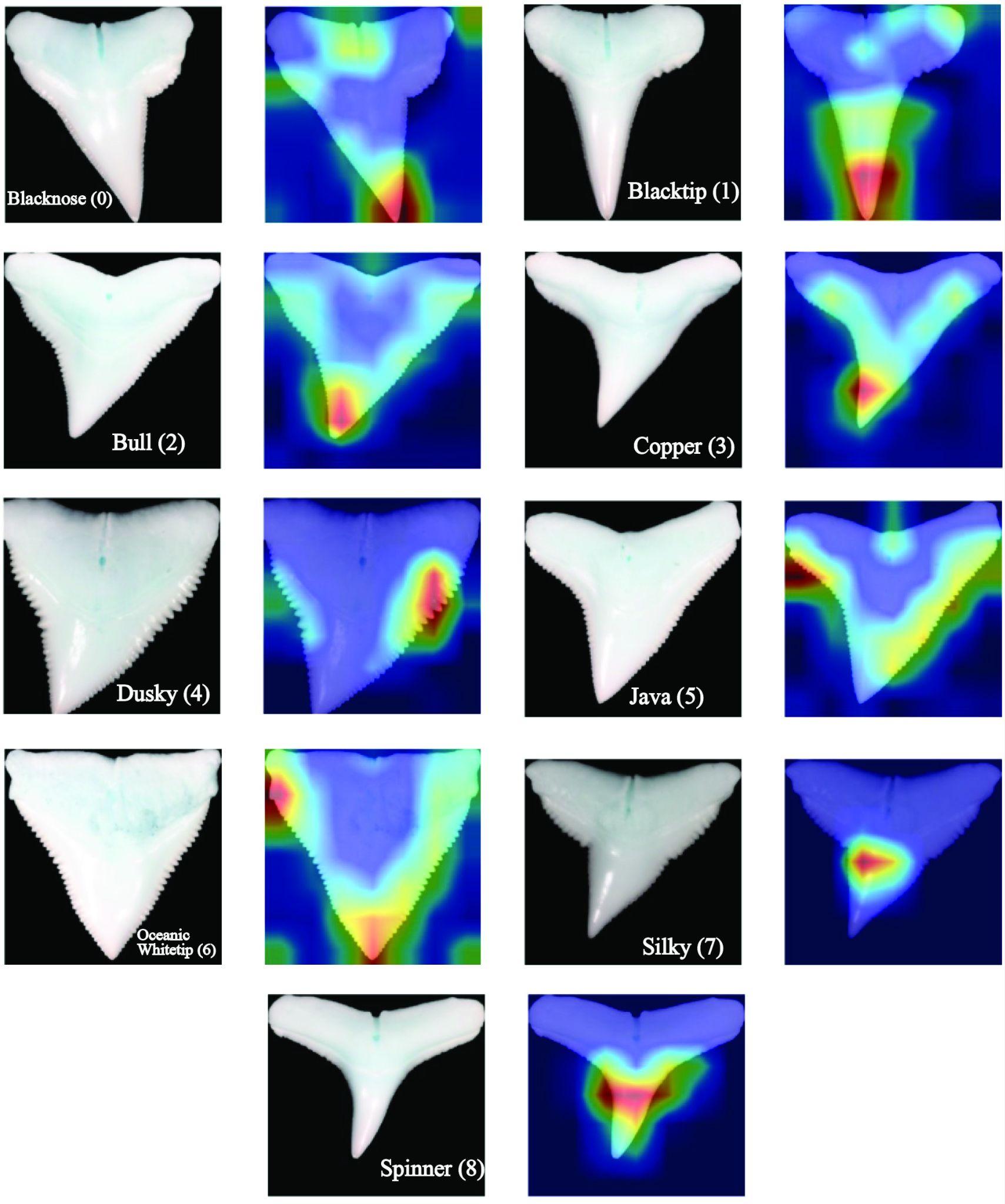


Figure 5: Gradient-CAM Feature mapping. Numbers correspond to the class label. Red color indicates high importance in classification, where blue indicates low importance.

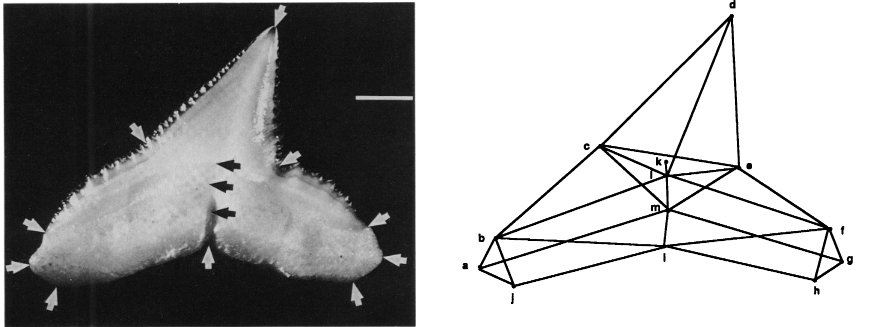


Figure 6: Figures 5 (left) and 6 (right) from Naylor & Marcus (1994) detailing the landmarks used for morphometric analysis.

|  |  |  |
| --- | --- | --- |
| Feature Map | Representative Landmarks  (in order of importance) | Corresponding Naylor & Marcus Landmark Letters |
| Blacknose (0) | Tip, nutrient groove | d, m |
| Blacktip (1) | Tip (general lower crown), nutrient groove | d, m |
| Bull (2) | Tip, shoulder serrations (distal and mesial), nutrient groove | d, e to f, m |
| Copper (3) | Distal bend in crown towards tip, gradient between root and enamel along shoulder (distal and mesial) | (not recognized) x 3 |
| Dusky (4) | Mesial notch/bend, distal notch/bend | c, e |
| Java (5) | Distal shoulder serrations, mesial bend, mesial edge, nutrient groove | (not recognized), c, b to d, m |
| Oceanic Whitetip (6) | Distal side cusplet, tip, mesial side cusplet | (not applicable), d, (not applicable) |
| Silky (7) | Distal notch\* | e \*(possibly on enamel) |
| Spinner (8) | Band across center crown, possibly width of crown | c to e |

Table 1: Representative landmarks and corresponding Naylor & Marcus landmark letters from figure 6.

Evidence from table 1 indicates that while the majority of representative landmarks agree with the Naylor & Marcus (1994) landmark system, there exist some new landmarks that could be incorporated into morphometric analysis of future research. For example, in the case of Copper (3), human visual analysis concludes that the feature map is highlighting the gradient/edge along the root and enamel along the shoulder on both distal and mesial sides. It is worth noting that the Naylor & Marcus system is rectified on the labial side, while our analysis is conducted on the lingual side. In our opinion, this side gives the CNN more features to work with e.g. the root shape. Incidentally, we hypothesized that the CNN would view features different to research that has been done, and, while this is the case in a few circumstances, a greater proportion of maps contained features documented by Naylor & Marcus (1994). Thus, the evidence further reinforces the Naylor & Marcus landmark system.

*Patch-based CNN*

The patch-based method has a multitude of use cases, primarily in medical imagery (Cruz-Roa et. al., 2014; Mousavi et. al., 2016; Xu et. al., 2015). Given the low classification accuracy of the patch-based method, there is a possibility of error in the processing or implementation of the model. We suspect that the parameters used might simply be better suited for the use in the specific medical uses rather than generalizable. Hou 2016 did mention how much of the implementation was “written explicitly for the specific input dimensions, as well as the format of having several patients each with their own image and distributing the patient level class appropriately.” For this reason, more testing is necessary for this specific use case. Additionally, the extensive resources necessary to train the patch-based model (24+ hours on AMD Ryzen 4000 series 9 CPU) inhibit our ability to effectively train and test. This model could be transferred to use GPU, but currently, it has not been compatible.

An interesting outcome of the patch-based modeling was that many of the patches contained high resolution images of the serrations of each tooth. Using only serrations, it would be interesting to determine if it were possible to predict species membership simply by the serrations. Using that information, would it then be possible to determine empirical relationships between different factors which make up the pattern of serration. Second, using the patch-based methodology, we could, given enough information, use the patchmaking to predict, given a tooth of a single species, jaw location, for we do have an overabundance of Dusky teeth. Utilizing extra oversight by the Computer Science department professors, we could better understand the nature of the low accuracy and model performance. Overall, an increased number of input training data images would undoubtedly perform with better classification, but much of this error was circumvented by the use of the pretrained VGG11 CNN, which was extremely effective at classifying given the inputs. These results, given limited input, are further evidence for the use of this process over morphometric analysis.

*Bag of Words Approach*

The bag of wordsmethod of classification boasted a high classification accuracy indicating its viability as an alternative to morphometric analyses for biological processes. The SVC approach has quite a few beneficial compromises over the CNN for our data. First, the SVC is capable of processing images of any resolution. This limits the need for much of the preprocessing outlined in the methods. Second, ORB is invariant to scale, as well as rotation, which further highlights ease of use. Finally, the bag of words + SVC method is much simpler to set up and use quickly. These features of the method outlined above make it exceptionally viable as a replacement or alternative to morphometric analysis.

Even with large amounts of intraspecies and tooth position variabilities in the dentitions, quantified in Smith (2015), both the SVC and CNN were able to effectively classify the teeth with accuracy greater than Smith (2015). The speed and ease at which the SVC operates makes a case for extensive usage in fields that require less expertise than tuning a CNN, while compromising some accuracy. The CNN achieves much higher classification accuracy and is able to give a much more in depth description of the features in question. Both methods carry weight as an alternative to morphometric and discriminant analysis as classification methods simply by reducing manpower hours needed to do the analysis. The Github repository is free and open source under the GNU General Public License v3.0 at <https://github.com/mikafur32/Shark-Species-Classification>, needing only images to run.

Conclusions

Despite the large proportion of *Carcharhinus* in the shark population and extensive research, the evolutionary history of *Carcharhinus* is not well understood. We found that SVC and CNN methods are both effective at classifying shark taxa. It was determined that the SVC was less effective at classifying the taxa, but is easier and faster to implement. The ease of implementation and lack of individualized output (at this time) makes it a good candidate for classifying between two species. The SVC approach would be most beneficial to distinguish between two species, particularly in the case of *Carcharhinus cerdale* vs *Carcharhinus porosus* (Castro, 2011). It was also observed that, in our analysis, species with visibly similar shapes are still more likely to be misclassified by the CNN. We hypothesize that this could be mitigated with additional training images. The difficulty in implementation and low classification accuracy dissuades further usage of the patch-based model in its current implementation, but the prospect of using the full resolution of our images gives it high desirability if it can be augmented. Our SVC, CNN, and patch-CNN analyses of tooth images are a stepping stone to revealing the true evolutionary history of the genus.

Appendix

|  |  |  |
| --- | --- | --- |
| Class label | Colloquial Name | Species Name |
| 0 | Blacknose | *Carcharhinus acronotus* |
| 1 | Blacktip | *Carcharhinus limbatus* |
| 2 | Bull | *Carcharhinus leucas* |
| 3 | Copper | *Carcharhinus brachyurus* |
| 4 | Dusky | *Carcharhinus obscurus* |
| 5 | Java | *Carcharhinus amboinensis* |
| 6 | Oceanic Whitetip | *Carcharhinus longimanus* |
| 7 | Silky | *Carcharhinus falciformis* |
| 8 | Spinner | *Carcharhinus brevipinna* |

References

Awad, Mariette & Khanna, Rahul. (2015). Support Vector Machines for Classification. 10.1007/978-1-4302-5990-9\_3.

Bondesan, A., Meneghel, M., & Sauro, U. (1992). Morphometric analysis of dolines. *International Journal of Speleology*, *21*(1), 1.

Bookstein, F., Schäfer, K., Prossinger, H., Seidler, H., Fieder, M., Stringer, C., ... & Marcus, L. F. (1999). Comparing frontal cranial profiles in archaic and modern Homo by morphometric analysis. *The Anatomical Record: An Official Publication of the American Association of Anatomists*, *257*(6), 217-224.

Castro, J. I. (2011). Resurrection of the name *Carcharhinus* cerdale, a species different from Carcharhinus porosus. Aqua International Journal of Ichthyology, 17(1), 1–10.

Cruz-Roa A, Basavanhally A, F. Gonz ́alez, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, and A. Madabhushi. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In Medical Imaging, 2014

C. Iorga and V. -E. Neagoe, "A Deep CNN Approach with Transfer Learning for Image Recognition," *2019 11th International Conference on Electronics, Computers and Artificial Intelligence (ECAI)*, 2019, pp. 1-6, doi: 10.1109/ECAI46879.2019.9042173.

Dosay-Akbulut, M. (2008). The phylogenetic relationship within the genus carcharhinus. *Comptes Rendus Biologies*, *331*(7), 500–509. <https://doi.org/10.1016/j.crvi.2008.04.001>

Garrick, J. A. F. (1982). Sharks of the genus Carcharhinus.

Harry, A. V., Morgan, J. A. T., Ovenden, J. R., Tobin, A. J., Welch, D. J., & Simpfendorfer, C. A. (2012). Comparison of the reproductive ecology of two sympatric blacktip sharks (Carcharhinus limbatus and Carcharhinus tilstoni) off north‐eastern Australia with species identification inferred from vertebral counts. *Journal of Fish Biology*, *81*(4), 1225-1233.

Hirshfield, A. N., & Rees Midgley Jr, A. (1978). Morphometric analysis of foilicular development in the rat. *Biology of reproduction*, *19*(3), 597-605.

Hou, Samaras, T. M. Kurc, Y. Gao, J. E. Davis and J. H. Saltz (2016), "Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification," IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, pp. 2424-2433, doi: 10.1109/CVPR.2016.266.

Kemp, N. E., & Westrin, S. K. (1979). Ultrastructure of calcified cartilage in the endoskeletal tesserae of sharks. *Journal of Morphology*, *160*(1), 75-101.

Long, J. A., Burrow, C. J., Ginter, M., Maisey, J. G., Trinajstic, K. M., Coates, M. I., Young, G. C., & Senden, T. J. (2015). First Shark from the late devonian (frasnian) gogo formation, Western Australia sheds new light on the development of tessellated calcified cartilage. *PLOS ONE*, *10*(5). https://doi.org/10.1371/journal.pone.0126066

Lowe, D. G. (1999, September). Object recognition from local scale-invariant features. In *Proceedings of the seventh IEEE international conference on computer vision* (Vol. 2, pp. 1150-1157). Ieee.

Joachims, T. (1998, April). Text categorization with support vector machines: Learning with many relevant features. In *European conference on machine learning* (pp. 137-142). Springer, Berlin, Heidelberg.

Mousavi H. S., Monga V., G. Rao, and A. U. Rao (2015). Automated discrimination of lower and higher grade gliomas based on histopathological image analysis. JPI,

Naylor, G.J.P. (1992), THE PHYLOGENETIC RELATIONSHIPS AMONG REQUIEM AND HAMMERHEAD SHARKS: INFERRING PHYLOGENY WHEN THOUSANDS OF EQUALLY MOST PARSIMONIOUS TREES RESULT. Cladistics, 8: 295-318.<https://doi.org/10.1111/j.1096-0031.1992.tb00073.x>

Naylor, G. J., & Marcus, L. F. (1994). Identifying isolated shark teeth of the genus Carcharhinus to species: relevance for tracking phyletic change through the fossil record. American Museum Novitates, (3109), 1–53.

Purdy, R. W., Schneider, V. P., Applegate, S. P., McLellan, J. H., Meyer, R. L., & Slaughter, B. H. (2001). The Neogene sharks, rays, and bony fishes from Lee Creek Mine, Aurora, North Carolina. Smithsonian Contributions to Paleobiology, 71-202(90).

Rublee, E., Rabaud, V., Konolige, K., & Bradski, G.R. (2011). ORB: An efficient alternative to SIFT or SURF. *2011 International Conference on Computer Vision*, 2564-2571.

Sandin, S. A., French, B. J., & Zgliczynski, B. J. (2022). Emerging insights on effects of sharks and other top predators on coral reefs. *Emerging Topics in Life Sciences*, *6*(1), 57–65. <https://doi.org/10.1042/etls20210238>

Seidel, R., Blumer, M., Pechriggl, E.-J., Lyons, K., Hall, B. K., Fratzl, P., Weaver, J. C., & Dean, M. N. (2017). Calcified cartilage or bone? Collagens in the tessellated endoskeletons of cartilaginous fish (sharks and rays). *Journal of Structural Biology*, *200*(1), 54–71. <https://doi.org/10.1016/j.jsb.2017.09.005>

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-cam: Visual explanations from deep networks via gradient-based localization. *2017 IEEE International Conference on Computer Vision (ICCV)*. <https://doi.org/10.1109/iccv.2017.74>

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Smith, V. (2015). *Species Discrimination in Carcharhinus Shark Teeth Using Eliptical Fourier Analysis* [Masters of Science, Tulane University]. ProQuest Dissertations Publishing. <https://www.proquest.com/docview/1689397480?pq-origsite=gscholar&fromopenview=true>

Smith, D. K. (1998). A morphometric analysis of Allosaurus. *Journal of Vertebrate Paleontology*, *18*(1), 126-142.

Soda, K. J. (2013). The Integration of Artificial Neural Networks and Geometric Morphometrics to Classify Teeth from Carcharhinus Species. Retrieved from <http://purl.flvc.org/fsu/fd/FSU_migr_etd-8640>

Xu Y., Jia Z., Y. Ai, F. Zhang, M. Lai, E. I. Chang, et al (2015). Deep convolutional activation features for large scale brain tumor histopathology image classification and segmentation.  
In ICASSP.

Zhang, Y., Jin, R. & Zhou, ZH. Understanding bag-of-words model: a statistical framework. *Int. J. Mach. Learn. & Cyber.* **1**, 43–52 (2010). https://doi.org/10.1007/s13042-010-0001-0