

# Identification of backed flakes in Discoid and RCLevallois

## Combining quantitative approaches to differentiate between backed products from discoidal and Levallois reduction sequences

Guillermo Bustos-Pérez<sup>1,2,3</sup>

Brad Gravina<sup>4,5</sup>

Michel Brenet<sup>5,6</sup>

Francesca Romagnoli<sup>1</sup>

<sup>1</sup>Universidad Autónoma de Madrid. Departamento de Prehistoria y Arqueología, Campus de Cantoblanco, 28049 Madrid, Spain

<sup>2</sup>Institut Català de Paleoecologia Humana i Evolució Social (IPHES), Zona Educacional 4, Campus Sesce-lades URV (Edifici W3), 43007 Tarragona, Spain

<sup>3</sup>Àrea de Prehistòria, Universitat Rovira i Virgili (URV), Avinguda de Catalunya 35, 43002 Tarragona, Spain

<sup>4</sup>Musée national de Préhistoire, MCC, 1 rue du Musée, 24260 Les Eyzies de Tayac, France

<sup>5</sup>UMR-5199 PACEA, Université de Bordeaux, Bâtiment B8, Allée Geoffroy Saint Hilaire, CS 50023, 33615 PESSAC CEDEX, France

<sup>6</sup>INRAP Grand Sud-Ouest, Centre mixte de recherches archéologiques, Domaine de Campagne, 242460 Campagne, France

Corresponding authors:

G.B.P. [guillermo.willbustos@mail.com](mailto:guillermo.willbustos@mail.com)

F.R. [f.romagnoli2@gmail.com](mailto:f.romagnoli2@gmail.com)

## Abstract

Backed flakes (core edge flakes and pseudo-Levallois points) represent special products of Middle Paleolithic centripetal flaking strategies. Their peculiarities are due to their roles as both a technological objective and in the management of core convexities to retain its geometric properties during reduction. In Middle Paleolithic contexts, these backed implements are commonly produced during Levallois and discoidal reduction sequences. Backed products from Levallois and discoidal reduction sequences often show common geometric and morphological features that complicate their attribution to one of these methods. This study examines the identification of experimentally produced discoidal and recurrent centripetal Levallois backed products (including all stages of reduction) based on their morphological features. 3D geometric morphometrics are employed to quantify morphological variability among the experimental sample. Dimensionality reduction through principal component analysis is combined with 11 machine learning models for the identification of knapping methods. A supported vector machine with polynomial kernel has been identified as the best model (with a general accuracy of 0.76 and an area under the curve [AUC] of 0.8). This indicates that combining geometric morphometrics, principal component analysis, and machine learning models succeeds in capturing the morphological differences of backed products according to the knapping method.

**Key words:** lithic analysis; Levallois; Discoid; Geometric Morphometrics; Machine Learning; Deep Learning

## 1. Introduction

The Middle Paleolithic in Western Europe is characterized by the diversification and increase of knapping methods resulting in flake-dominated assemblages (Delagnes and Meignen, 2006; Kuhn, 2013). Discoidal and the recurrent centripetal Levallois are two of the most common flake production systems during this period. Following Boëda (1995a, 1994, 1993), there are six technological criteria that define discoidal debitage:

1. The volume of the core is conceived as two oblique asymmetric convex surfaces delimited by an intersection plane;
2. These two surfaces are not hierarchical being possible to alternate the roles of percussion and exploitation surfaces;
3. The peripheral convexity of the debitage surface is managed to control lateral and distal extractions thus allowing for a degree of predetermination;
4. Surfaces of the striking planes are oriented in a way that the core edge is perpendicular to the predetermined products;
5. The fracture planes are secant;
6. The technique employed is the direct percussion with hard hammer.

In addition, according to Boëda (1994, 1993) six characteristics define the Levallois knapping strategy:

1. The volume of the core is conceived in two convex asymmetric surfaces;
2. These two surfaces are hierarchical and are not interchangeable. They maintain their role of striking and debitage (or exploitation) surface respectively along the whole reduction process;
3. The distal and lateral convexities of the debitage surface are maintained to obtain predetermined flakes;
4. The fracture plane of the predetermined products is parallel to the intersection between both surfaces;
5. The striking platform is perpendicular to the overhang (the core edge, at the intersection between the two core surfaces);
6. The technique employed during the knapping process is the direct percussion with hard hammer.

Depending on the organization of the debitage surface Levallois cores are usually classified into preferential method (where a single predetermined Levallois flake is obtained from the debitage surface) or recurrent methods (where several predetermined flakes are produced from the debitage surface) with removals being either unidirectional, bidirectional or centripetal (Boëda, 1995a; Delagnes, 1995; Delagnes and Meignen, 2006).

Both knapping methods involve the removal of backed products that usually comprise two categories: core edge flakes (*eclat débordant*) and pseudo-Levallois points.

Core edge flakes / *eclat débordant* (Beyries and Boëda, 1983; Boëda, 1993; Boëda et al., 1990) are technical backed knives that have a cutting edge opposite and parallel (or sub-parallel) to an abrupt margin (a back that usually has an angle close to 90°). This back commonly results from the removal of one of the lateral edges of the core and can be plain, retain the scars from previous removals, be cortical, or a combination of these attributes. Core edge flakes are also divided into two categories: “classic core edge flakes” and “core edge flakes with a limited back.” “Classic core edge flakes” (Beyries and Boëda, 1983; Boëda, 1993; Boëda et al., 1990), which are sometimes referred to as “core edge flakes with a non-limited back” / “*eclat débordant*

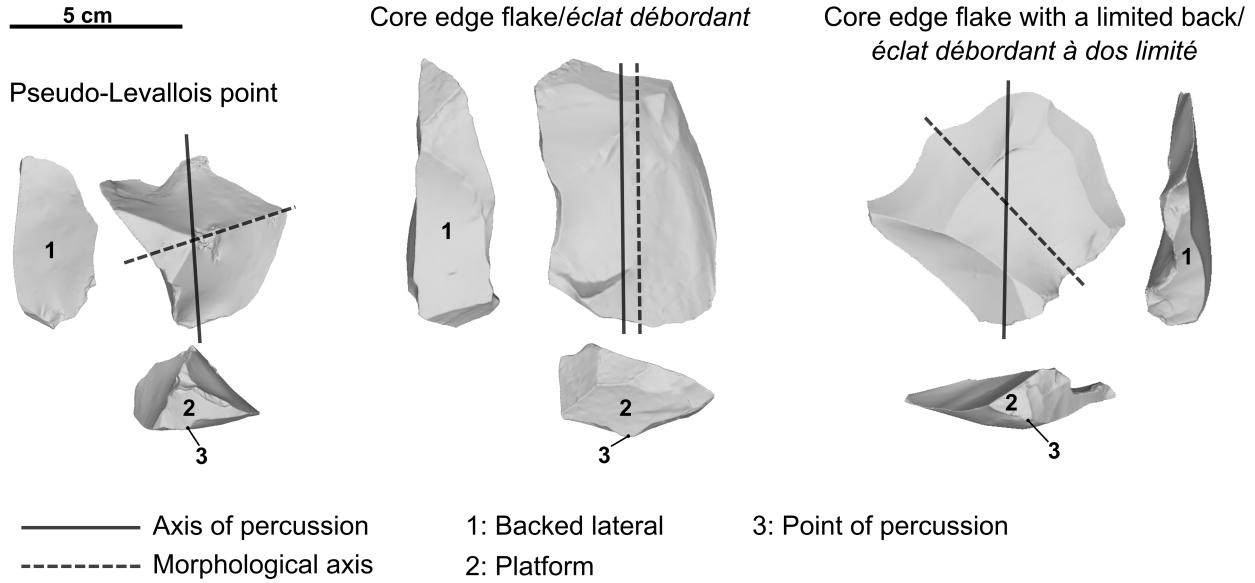


Figure 1: Examples of backed flakes categories with key features discussed in the present research)

*à dos non limité*” (Duran, 2005; Duran and Soler, 2006), have a morphological axis more or less similar to the axis of percussion. “Core edge flakes with a limited back”/ “éclat débordant à dos limité” a offset axis of symmetry in relation to the axis of percussion (Meignen, 1996; Meignen, 1993; Pasty et al., 2004). This orientation often leads to the back not being parallel to nor spanning the entire length of the sharp edge or the percussion axis (Slimak, 2003).

Pseudo-Levallois points (Boëda, 1993; Boëda et al., 1990; Bordes, 1961, 1953; Slimak, 2003) are backed products where the edge opposite to the back has a triangular morphology. This triangular morphology is usually the result of the convergence of two or more scars. As with core edge flakes, the back usually results from the removal of one of the lateral edges of the core and can be plain, retain the scars from previous removals, or more rarely be cortical or a combination of these traits. Both pseudo-Levallois points and core edge flakes with a limited back share a symmetry offset from the axis of percussion but are clearly differentiable due to their morphology. The present study includes the three categories defined above as backed products.

Depending on the knapping method, different roles in Levallois recurrent centripetal and discoidal debitage are attributed to core edge flakes and pseudo-Levallois points. Boëda et al. (1990) focus on the role of core edge flakes and cortically backed flakes for maintaining the lateral convexities throughout Levallois recurrent centripetal reduction. Similarly, pseudo-Levallois points contribute to maintaining the lateral and distal convexities between different series of removals (Boëda et al., 1990).

Focusing on the variability of discoidal debitage, Slimak (2003) noted that pseudo-Levallois points are short products that induce a limited lowering of the core overhang (the intersection between the striking and debitage surfaces). In contrast, core edge flakes can result from several distinct production objectives. Expanding on the roles of pseudo-Levallois points and core edge flakes within discoidal debitage, Locht (2003) demonstrated the systematic production of both products at the site of Beauvais. This indicates that at Beauvais, core edge flakes and pseudo-Levallois points were the main predetermining/predetermined products (Locht, 2003).

An additional added value of core edge flakes and pseudo-Levallois points is their frequent transport by Paleolithic groups. Turq et al. (2013) described the widespread import and export of lithic artifacts during the Middle Paleolithic. Examples of the transport of pseudo-Levallois points from discoidal production sequences can be observed at Combemenu, La Mouline, Les Fieux (Brenet, 2013, 2012; Brenet and Cretin, 2008; Folgado and Brenet, 2010; Turq et al., 2013), and the open-air site of Bout des Vergnes (Courbin et



Figure 2: Middle Paleolithic sites cited in the text showing examples of transport of backed products from discoidal or Levallois recurrent centripetal sequences (base map obtained from <https://maps-for-free.com/>)

al., 2020), while the transport of core edge flakes (into and out of the site) is also clearly observed at la Grotte Vaufrey (Geneste, 1988), Teixoneres cave (Bustos-Pérez et al., 2017; Picin et al., 2020), Amalda Cave (Rios-Garaizar, 2010), Grotta del Cavallo (F. Romagnoli et al., 2016), l'Arbreda (Duran and Soler, 2006) and at Site N of Maastricht-Belvédère (Roebroeks et al., 1992). Transported backed pieces have also been clearly identified at Abric Romaní in Spain within both Levallois and discoidal production methods (Martín-Viveros et al., 2020; Francesca Romagnoli et al., 2016). The identification of knapping methods among these specific products can help to identify diachronic and synchronic changes in lithic production, selection, and transport as a reflection of hominin technological organization and adaptive strategies (Binford, 1979; Shott, 2018)

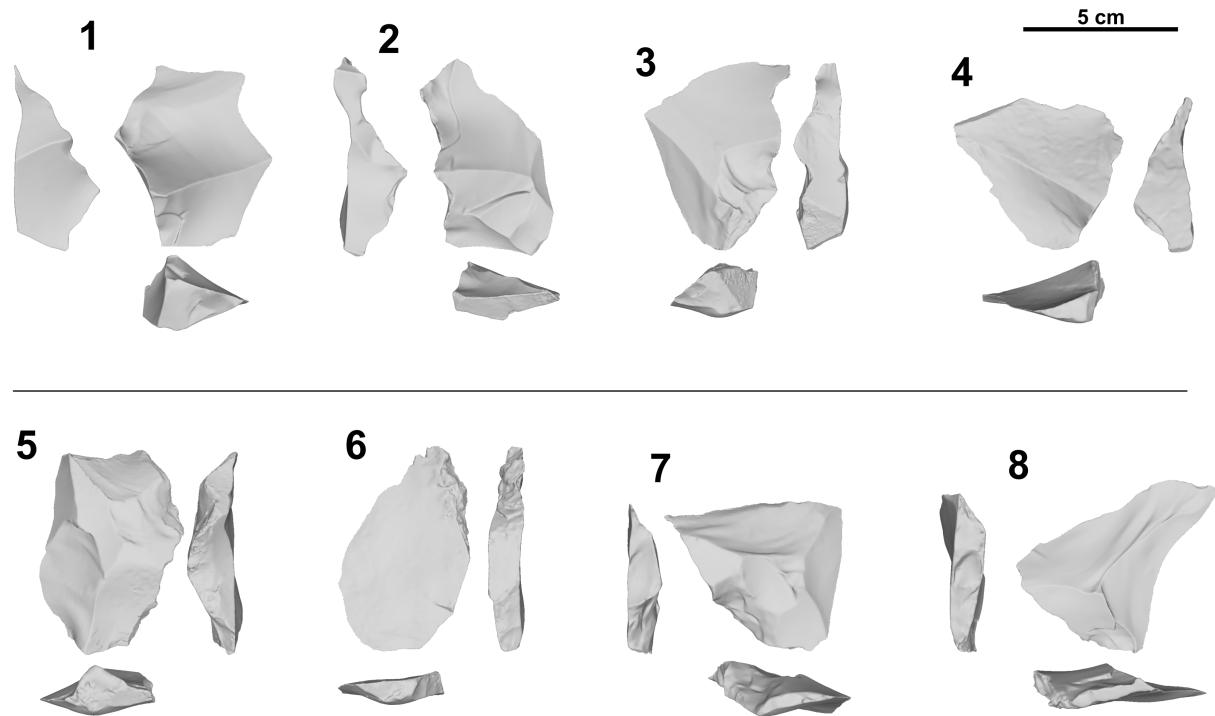


Figure 3: Backed products from the experimental sample: core edge flakes (1–2) and pseudo-Levallois points (3–4) from the Discoidal knapping method. Core edge flakes (5–6) and pseudo-Levallois points (7–8) from the Levallois recurrent centripetal method

The attribution of backed pieces to either discoidal or recurrent centripetal Levallois reduction can, however, be problematic. For example, Mourre (2003) indicates that a key aspect for the identification of Levallois core edge flakes is the direction of the debitage axis, which is parallel to the intersection plane of the two core surfaces while the fracture plane is secant. Slimak (1998) showed core edge flakes from discoidal reduction to equally have fracture planes parallel to the intersection between the debitage surface and striking platforms although not as parallel as in Levallois debitage. Delpiano et al. (2021) demonstrated a tendency of Levallois products to be more elongated with thinner and sub-parallel edges, whereas discoidal backed products show a higher variation in the minimum and maximum thickness of the back. Previous studies (Archer et al., 2021; González-Molina et al., 2020) addressed the differentiation between discoidal and recurrent centripetal Levallois products in general terms were all products are considered. While this approach is highly effective, the differentiation between backed products of discoidal and recurrent centripetal Levallois sequences is not sufficiently addressed. Given the special technological role of backed products in core management and production, their specific techno-functional properties (Delpiano et al., 2021), the frequency in which they appear in the archaeological record, and their common transport as part of hominin toolkits, a more systematic approach to their accurate differentiation represents an important advancement in describing Middle Palaeolithic lithic assemblages.

This raises the issue as to the extent to which discoidal and Levallois recurrent centripetal core edge flakes and pseudo-Levallois points can be differentiated based on their morphological features. This issue is relevant to lithic studies because it affects the technological analysis of a stone tool assemblage and the evolutionary interpretation of knapping concepts over time. Here we address this issue through experimental archaeology and a multi-level statistical approach. We reproduced classic bifacial discoidal and recurrent centripetal Levallois reduction sequences to obtain a collection of backed products. We produced 3D scans of lithic artifacts and employed geometric morphometrics to quantify the morphological variability of the experimental sample and the cores were refit. Dimensionality reduction through principal component analysis (PCA) was carried out on a set of coordinates, and 11 machine learning models were tested to obtain classification accuracy and variable importance. Geometric morphometrics and Machine Learning models make it possible to directly test technological classifications of lithic and features usually employed to discriminate between both methods.

## 2 Methods

### 2.1 Experimental assemblage

The analyzed experimental assemblage derives from the replication of nine discrete knapping sequences. Seven cores were knapped in Bergerac chert (Fernandes et al., 2012), and two cores were knapped in Miocene chert from South of Madrid (Bustillo et al., 2012; Bustillo and Pérez-Jiménez, 2005). Five cores were knapped following the discoidal “*sensu stricto*” method, which corresponds highly to Boëda’s original technological definition of the knapping system (Boëda, 1995b, 1994, 1993), and five experimental cores were knapped following the Levallois recurrent centripetal system (Boëda, 1995a, 1994, 1993; Lenoir and Turq, 1995). A total of 139 unretouched backed flakes (independent of the type of termination) were obtained: 70 from the discoidal reduction sequences and 69 from the Levallois reduction sequences (Figure 1). In the case of the Levallois recurrent centripetal cores, backed products from both debitage and striking surfaces were included.

The following code loads data and packages necessary for the development of the present research.

```
# List of packages
list.of.packages <- c("tidyverse", "caret", "ranger")

# Load packages
lapply(list.of.packages, library, character.only = TRUE)

## [[1]]
## [1] "knitr"      "ranger"      "caret"       "lattice"     "forcats"     "stringr"     "dplyr"
## [8] "purrr"       "readr"       "tidyR"       "tibble"      "ggplot2"     "tidyverse"   "stats"
## [15] "graphics"    "grDevices"   "utils"       "datasets"    "methods"     "base"
##
## [[2]]
## [1] "knitr"      "ranger"      "caret"       "lattice"     "forcats"     "stringr"     "dplyr"
## [8] "purrr"       "readr"       "tidyR"       "tibble"      "ggplot2"     "tidyverse"   "stats"
## [15] "graphics"    "grDevices"   "utils"       "datasets"    "methods"     "base"
##
## [[3]]
## [1] "knitr"      "ranger"      "caret"       "lattice"     "forcats"     "stringr"     "dplyr"
## [8] "purrr"       "readr"       "tidyR"       "tibble"      "ggplot2"     "tidyverse"   "stats"
## [15] "graphics"    "grDevices"   "utils"       "datasets"    "methods"     "base"
```

```

rm(list.of.packages)

# Loading landmarks coordinates
load("Data/Flakes LM rotated.RData")

# Loading manual attributes
Att <- read.csv("Data/Attributes data.csv")

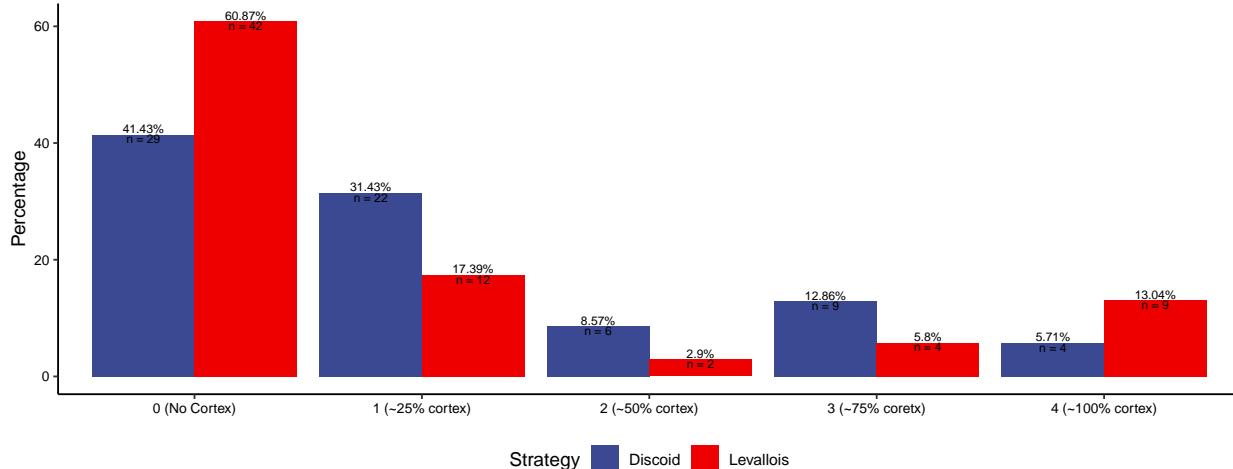
```

The Levallois recurrent centripetal experimental assemblage is clearly dominated by non-cortical backed flakes ( $n = 42$ ; 60.87%). This is expected since one of the roles of core edge flakes and pseudo-Levallois points in Levallois recurrent centripetal methods is the management of convexities on subsequent exploitation sequences (Boëda, 1994, 1993; Boëda et al., 1990). Thus, although backed flakes can be present in the initial decortication phases ( $n = 9$ ; 13.04%), the subsequent exploitation of the core will result in non-cortical flakes. Non-cortical backed flakes are also the majority class of the experimental Discoid assemblage although this predominance is attenuated ( $n = 29$ ; 41.43%). However, along with flakes with nearly 25% of the dorsal surface covered with cortex, they make up the majority of the discoidal backed flakes of the assemblage ( $n = 51$ ; 72.86%). This reduction in the predominance of non-cortical flakes is also expected in discoidal methods given the organization of both debitage surfaces, the nature of the surface convexities, and the fracture plane. In discoidal cores, the interchangeable surfaces usually have a higher apical convexity than Levallois cores. Additionally, the angle and removal of flakes cover a smaller portion of the respective surface than in a Levallois core. Thus, it is expected that as reduction continues, some products will retain a certain amount of cortex.

```

# Cortex & method
Att %>% group_by(Strategy) %>%
  count(CORTEX) %>%
  mutate(Percentage = round(n/sum(n)*100, 2)) %>%
  ggplot(aes(CORTEX, Percentage, fill = Strategy)) +
  geom_col(position = "dodge") +
  ggsci::scale_fill_aaas() +
  xlab(NULL) +
  geom_text(aes(label = paste0(Percentage, "%")),
            vjust = -0.2, size = 2.5,
            position = position_dodge(.9)) +
  geom_text(aes(label = paste("n =", n)),
            vjust = "top", size = 2.5,
            position = position_dodge(.9)) +
  theme_classic() +
  theme(
    legend.position = "bottom",
    axis.text = element_text(color = "black", size = 8))

```



## 2.2 Data acquisition

All flakes were scanned with an Academia 20 structured light surface scanner (Creaform 3D) at a 0.2 mm resolution. Flakes were scanned in two parts and automatically aligned (or manually aligned in case automatic alignment failed) and exported in STL formats. Cloudcompare 2.11.3 (<https://www.danielgm.net/cc/>) free software was employed to perform additional cleaning, mesh sampling, surface reconstruction, and transformation into PLY files. Finally, all files were decimated to a quality of 50,000 faces using the Rvcg v.0.21 R package (Schlager, 2017).

The protocol for the digitalization of landmarks on flakes was based on previous studies (Archer et al., 2021, 2018). This included the positioning of a total of 3 fixed landmarks, 85 curve semi-landmarks, and 420 surface semi-landmarks (Bookstein, 1997a, 1997b; Gunz et al., 2005; Gunz and Mitteroecker, 2013; Mitteroecker and Gunz, 2009). This makes for a total of 508 landmarks and semi-landmarks. The three fixed landmarks correspond to both laterals of the platform width, and the percussion point. The 85 curve semi-landmarks correspond to the internal and exterior curve outlines of the platform (15 semi-landmarks each) and the edge of the flake (55 semi-landmarks), and the 60 surface semi-landmarks correspond to the platform surface. The dorsal and ventral surfaces are defined by 180 semi-landmarks each. The workflow for digitalizing the landmarks and semi-landmarks included the creation of a template/atlas on an arbitrary selected flake. After this, the landmarks and semi-landmarks were positioned in each specimen and were relaxed to minimize bending energy (Bookstein, 1997a, 1997b). The entire workflow of landmark and semi-landmarks digitalization and relaxation to minimize bending energy was done in Viewbox version 4.1.0.12 (<http://www.dhal.com/viewbox.htm>), and resulting point coordinates were exported into .xlsx files.

Procrustes superimposition (Kendall, 1984; Mitteroecker and Gunz, 2009; O'Higgins, 2000) was performed using the package "Morpho" v.2.9 (Schlager, 2017) on RStudio IDE (R. C. Team, 2019; Rs. Team, 2019). After performing Procrustes superimposition and obtaining a new set of coordinates, PCA was performed to reduce the dimensionality of the data (James et al., 2013; Pearson, 1901). There are multiple reasons to use dimensionality reduction when dealing with high dimensional data on classification, including to avoid having more predictors than observations ( $p > n$ ), to avoid the collinearity of predictors, to reduce the dimensions of the feature space, and to avoid overfitting due to an excessive number of degrees of freedom (simple structure with lower number of variables). PCA achieves dimensionality reduction by identifying the linear combinations that best represent the predictors in an unsupervised manner. The principal components (PCs) of a PCA are aimed to capture as high a variance as possible of the complete data (James et al., 2013), and PCs that capture a higher variance do not necessarily need to be the best for classification. For the present work, PCs that represent 95% of the variance were selected as predictors for training the machine learning models. The threshold of 95% of the variance was arbitrarily selected since it balances retaining most of the dataset variance on a reduced number of variables. The identification of best PCs for classification was automatically done by the machine learning models using the caret v.6.0.92 package (Kuhn, 2008).

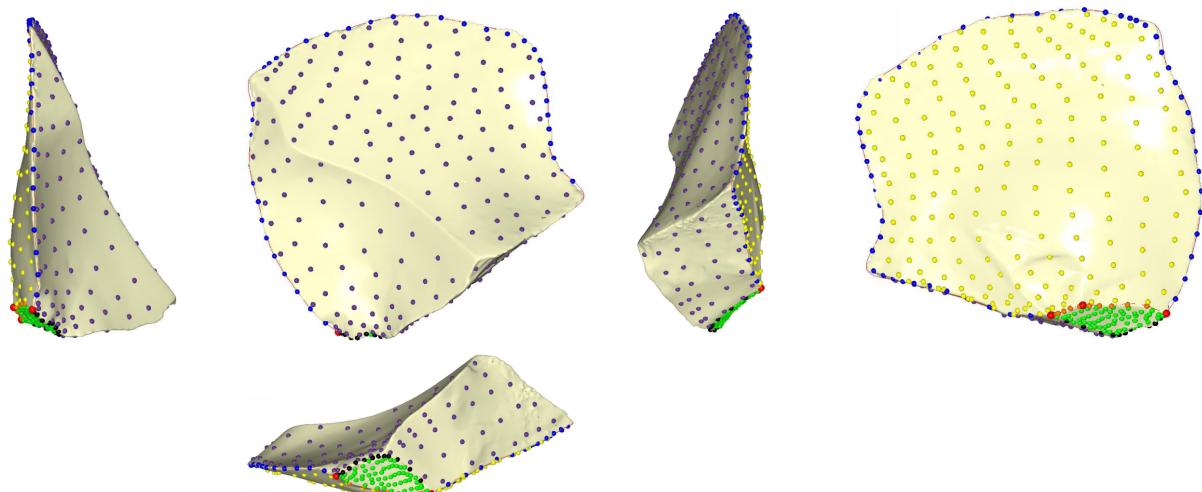
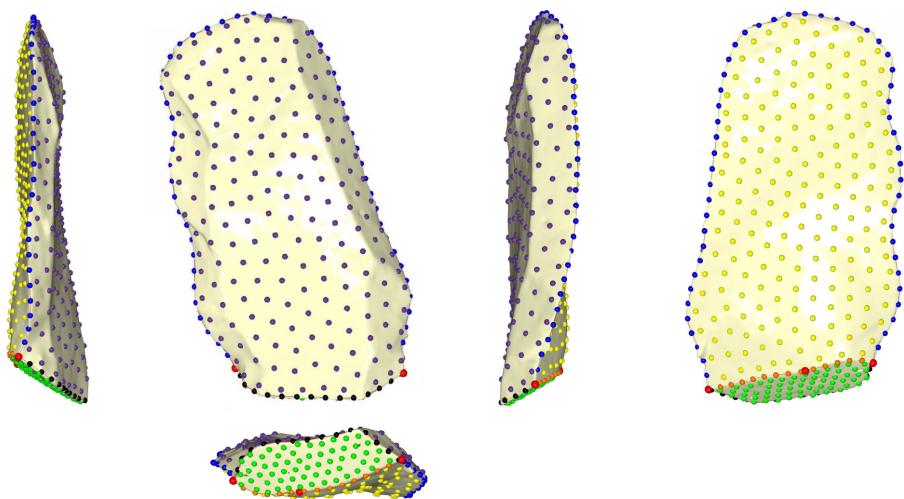


Figure 4: Template/atlas on an arbitrary selected flake with the defined landmarks (in red), curves, and surfaces. Bottom: Landmark positioning after sliding to minimize bending energy on a pseudo-Levallois point from a discoidal reduction sequence.

A drawback exists in the interpretability of PCA as a result of its nature of reducing dimensionality through the identification of linear combinations of variables. Because of this, interpreting what morphological features are being captured by a PC can be difficult. An option for interpreting PC is to use manual measurements as predictive variables on multiple linear regressions to predict PC values. In addition to geometric morphometrics, the following attributes were recorded for each of the flakes using the E5 software ([McPherron, 2019](#)).

- **Technological length:** measured in mm along the axis perpendicular to the striking platform.
- **Technological width:** measured in mm along the axis perpendicular to the technological length.
- **Maximum thickness** of the flake measured in mm.
- **External platform angle (EPA):** measured in degrees with a manual goniometer.
- **Internal platform angle (EPA):** measured in degrees with a manual goniometer.
- **Relative amount of cortex on the dorsal face:** recorded according to its extension on the dorsal surface of the flake, with categories as follows: 0 (no cortex), 1 (nearly 25% covered by cortex), 2 (nearly 50% covered by cortex), 3 (nearly 75% covered by cortex), and 4 (nearly the entire surface covered by cortex). This variable was employed to evaluate the distribution of cortex proportions among the experimental assemblage.
- **Weight:** measured to a precision of 0.01 g.

These measures served to generate the following indices:

- **Elongation index:** length divided by width.
- **Crenation index:** result of dividing either width or length (the one with the lowest value) between maximum thickness.
- **Width to thickness ratio:** flake width divided by maximum thickness.

These measures are not intended as inputs for the Machine Learning models, but to explore the meaning of the Principal Components through multiple linear regressions.

### 2.3 Machine learning models and evaluation

The following 11 machine learning models have been tested for differentiating between backed flakes extracted from the two surfaces of the core within each knapping method:

- **Linear discriminant analysis (LDA):** reduces dimensionality aiming to maximize the separation between classes while decision boundaries divide the predictor range into regions ([Fisher, 1936; James et al., 2013](#)).
- **K-nearest neighbor (KNN):** classifies cases by assigning the class of similar known cases. The “k” in KNN references the number of cases (neighbors) to consider when assigning a class, and it must be found by testing different values. Given that KNN uses distance metrics to compute nearest neighbors and that each variable is in different scales, it is necessary to scale and center the data prior to fitting the model ([Cover and Hart, 1967; Lantz, 2019](#)).

- **Logistic regression:** essentially adapts continuous regression predictions to categorical outcomes ([Cramer, 2004](#); [Walker and Duncan, 1967](#)).
- **Decision tree with C5.0 algorithm:** is an improvement on decision trees for classification ([Quinlan, 2014](#); [Quinlan, 1996](#)).
- **Random fores:** is made of decision trees. Each tree is grown from a random sample of the data and variables, allowing for each tree to grow differently and to better reflect the complexity of the data ([Breiman, 2001](#)).
- **Gradiant Boosting Machine:** ([Greenwell et al., 2019](#); [Ridgeway, 2007](#)) implements gradient boosted ([Friedman, 2002, 2001](#)), allowing the detection of learning deficiencies and increases model accuracy.
- **Supported vector machine (SVM):** fits hyperplanes into a multidimensional space with the objective of creating homogeneous partitions ([Cortes and Vapnik, 1995](#); [Frey and Slate, 1991](#)). The present study tests SVM with linear, radial, and polynomial kernels.
- **Artificial neural network (ANN):** with multi-layer perception, uses a series of hidden layers and error backpropagation for model training ([Rumelhart et al., 1986](#)).
- **Naïve Bayes:** computes class probabilities using Bayes' rule ([Weihs et al., 2005](#)).

All models are evaluated using  $10 \times 50$  k-fold cross validation (10 folds and 50 cycles), providing measures of accuracy. Using a 10-fold division, each fold will have 14 data points (with the exception of the last fold, which will have 13 data points). Each fold serves subsequently as test set for a trained model. Although computationally more expensive, this guarantees that all data points will serve as test sets. The 50 cycles provide a random shuffling of the dataset prior to fold division, thus ensuring that the composition of the folds varies in each cycle and it does not play a significant role in the evaluation of the models.

Machine Learning models commonly use a 0.5 classification threshold to assign categories. However, classification thresholds can be modified to balance the ability of model to detect true positives and avoid false positives which are respectively referred as sensitivity and specificity (this problem is exemplified in the following figure). The receiver operating characteristic (ROC) curve is employed to evaluate the ratio of detected true positives while avoiding false positives ([Bradley, 1997](#); [Spackman, 1989](#)). The ROC curve allows visually analyzing model performance and calculating the AUC, which ranges from 1 (perfect classifier) to 0.5 (random classifier). AUC ranges of values are usually interpreted as follows: 1 to 0.9: outstanding; 0.9 to 0.8: excellent/good; 0.8 to 0.7: acceptable/fair; 0.7 to 0.6: poor; and 0.6 to 0.5: no discrimination ([Lantz, 2019](#)). When analyzing lithic assemblages, the use of thresholds to guarantee true positives and avoid false positives is of special interest. The use of decision thresholds and derived measures of accuracy (ROC curve and AUC) can be especially useful in lithic analysis since it is expected that products from initial reduction stages are morphologically similar independent of the knapping method. It is expected that these products show a higher mixture between methods and have lower probability values. The use of thresholds better indicates the accuracy of a model taking into account these probability values. The use of thresholds better indicates the accuracy of a model since it takes into account these probability values.

Statistical analysis was carried out using R version R version 4.1.1 in IDE RStudio version 2021.09.0 ([R. C. Team, 2019](#); [Rs. Team, 2019](#)). The management of the data and graphs was done using the tidyverse v.1.3.1 package ([Wickham et al., 2019](#)). The training of LDA and KNN was done with MASS v.7.3.57 ([Wright and Ziegler, 2017](#)). The training of SVM was done using the e1071 v.1.7.9 package ([Karatzoglou et al., 2006, 2004](#)). The RSNNS v.0.4.14 ([Bergmeir and Benítez, 2012](#)) package was employed to train multi-layer ANN with backpropagation. The klaR v.1.7.0 package was employed to train the naïve Bayes classifier ([Weihs et al., 2005](#)). The k-fold cross validation of all models, precision metrics, and confusion matrix were obtained using the caret v.6.0.92 package ([Kuhn, 2008](#)). Machine learning models also provide insights into variable importance for classification. The caret package was employed to extract variable importance after each

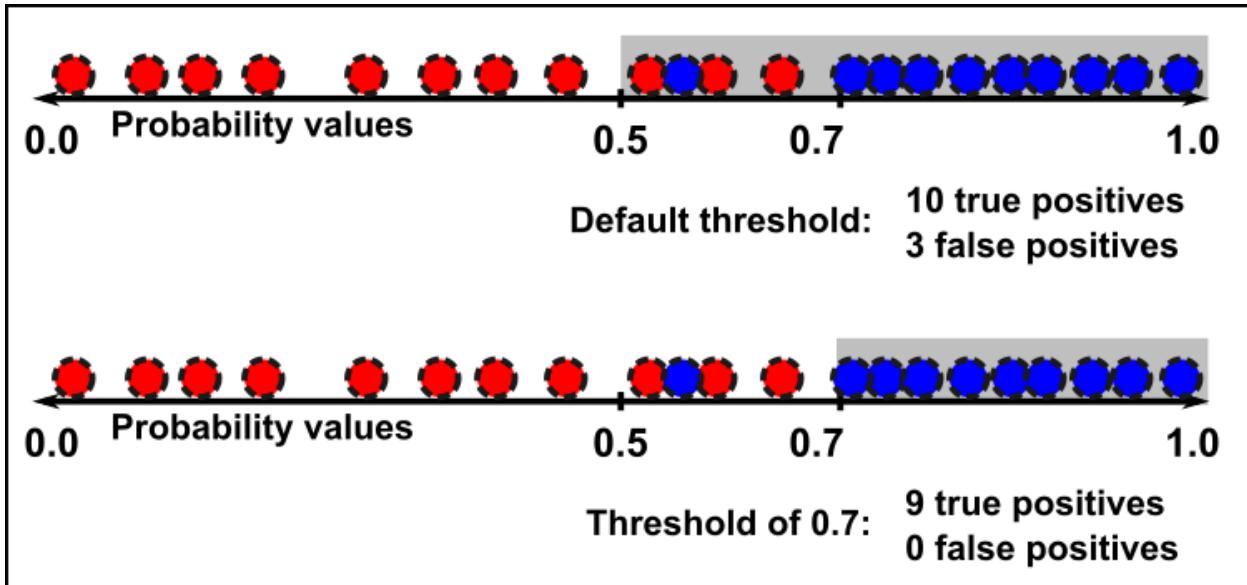


Figure 5: Examples of the use of different thresholds for classification. The upper example uses a default threshold of 0.5 for classification. The lower example uses a 0.7 threshold avoiding three false positives at the expense of one true positive

k-fold cross validation. Package pROC v.1.18.0 is employed to obtain ROC curve and AUC data.

#### 2.4 Performance of procrustes, PCA and model training

The following line of code performs procrustes alignment and superimposition using the **Morpho** package (Schlager, 2017). Aligned coordinates are extracted and stored as a data frame named **LM.DF**.

```
# Load landmarks
load("Data/Flakes_LM.RData")
# Procrustes alignment
proc <- Morpho::ProcGPA(Flakes_LM,
                        CSinit = TRUE,
                        silent = FALSE)

# Extract coordinates
Proc.Rot <- proc$rotated
LM.DF <- data.frame(matrix(Proc.Rot, nrow = length(filenames), byrow = TRUE))
```

The following line of code performs Principal Components Analysis (PCA) on the set of aligned coordinates stored in the LM.DF data frame. Summary provides proportion and cumulative proportion of variance explained by the 25 first principal components which add up to 95% of variance.

```
# PCA on coordinates
pca <- prcomp(LM.DF, scale. = TRUE)
summary(pca)$importance[1:3, 1:25]
```

	PC1	PC2	PC3	PC4	PC5	PC6	PC7
##							
## Standard deviation	18.05288	16.62783	12.83087	10.83128	10.42072	8.299316	7.73039

```

## Proportion of Variance  0.21385  0.18142  0.10803  0.07698  0.07125 0.045200 0.03921
## Cumulative Proportion  0.21385  0.39527  0.50330  0.58028  0.65153 0.696730 0.73594
##                               PC8      PC9      PC10     PC11     PC12     PC13     PC14
## Standard deviation    7.439897 6.710911 6.173021 5.368746 4.773021 4.562145 4.44228
## Proportion of Variance 0.036320 0.029550 0.025000 0.018910 0.014950 0.013660 0.01295
## Cumulative Proportion  0.772260 0.801810 0.826810 0.845730 0.860670 0.874330 0.88728
##                               PC15     PC16     PC17     PC18     PC19     PC20     PC21     PC22
## Standard deviation    3.803028 3.750857 3.54599 3.23142 2.956198 2.70170 2.649925 2.516422
## Proportion of Variance 0.009490 0.009230 0.00825 0.00685 0.005730 0.00479 0.004610 0.004160
## Cumulative Proportion  0.896770 0.906000 0.91425 0.92110 0.926840 0.93163 0.936240 0.940390
##                               PC23     PC24     PC25
## Standard deviation    2.466077 2.38239 2.327281
## Proportion of Variance 0.003990 0.00372 0.003550
## Cumulative Proportion  0.944380 0.94811 0.951660

```

Once PCA is performed, values of each of the 25 first PC's can be extracted for each backed flake. Additionally it is necessary to store the PC values along with each backed flake ID. Knapping method is documented for each of the experimental cores and can be added using the `case_when()` function.

```

# Store PCA values in a dataframe and add ID's
PCA_Coord <- as.data.frame(pca$x)
PCA_Coord$ID <- filenames
PCA_Coord$Core <- str_sub(PCA_Coord$ID, end = 2)

# Set the core to which they belong and strategy
PCA_Coord <- PCA_Coord %>% mutate(
  Strategy = case_when(Core == "B2" | Core == "B3" |
    Core == "B4" | Core == "B5" | Core == "B6" ~ "Discoid",
    Core == "B7" | Core == "B8" | Core == "B9" | Core == "Le" ~ "Levallois" ))

# Set strategy as factor or varImp will not work
PCA_Coord$Strategy <- factor(PCA_Coord$Strategy)

```

The resulting `PCA_Coord` data frame haves 25 numeric variables (values of the 25 first PC) along with artifact ID and associated core knapping strategy. This allows to train the models to predict knapping strategy "Strategy" based on the values of the 25 first PC. Here, the training of the models is done in three steps:

- 1) Set the formula.
- 2) Set the training control and validation method.
- 3) Train the models using the formula and validation method.

```

# Set formula
frmla <- as.formula(
  paste("Strategy", paste(colnames(PCA_Coord[,1:25]), collapse = " + "), sep = " ~"))

# Set cross validation
trControl <- trainControl(method = "repeatedcv",
                           verboseIter = TRUE,
                           number = 10,
                           repeats = 50,

```

```

                savePredictions = "final",
                classProbs = TRUE)

# LDA model
set.seed(123)
fit.LDA <- caret::train(frmla,
                        PCA_Coord,
                        method = "lda",
                        preProc = c("center", "scale"),
                        trControl = trControl)

# KNN model
set.seed(123)
KNN.model <- caret::train(
  frmla,
  PCA_Coord,
  method = "knn",
  preProc = c("center", "scale"),
  trControl = trControl,
  tuneGrid = expand.grid(k = seq(1, 15, 1)))
)

# Logistic regression model
set.seed(123)
logmod <- caret::train(frmla,
                        PCA_Coord,
                        method = "glm",
                        family = "binomial",
                        preProc = c("center", "scale"),
                        trControl = trControl)

# SVM linear
set.seed(123)
SVM_Linear <- train(frmla,
                      PCA_Coord,
                      method = "svmLinear",
                      preProcess = c("center", "scale"),
                      trControl = trControl,
                      tuneGrid = expand.grid(C = seq(0.01, 3, length = 20)),
                      metric = "Accuracy",
                      importance = 'impurity')

# SVM Radial
set.seed(123)
SVM_Radial <- train(frmla,
                      PCA_Coord,
                      method = "svmRadial",
                      preProcess = c("center", "scale"),
                      trControl = trControl,
                      tuneGrid =
                        expand.grid(C = seq(0.01, 3, length = 20),
                                    sigma = seq(0.0001, 1, length = 20)),
                      metric = "Accuracy",

```

```

importance = 'impurity')

# SVM Poly
set.seed(123)
SVM_Poly <- train(frmla,
                    PCA_Coord,
                    method = "svmPoly",
                    preProcess = c("center", "scale"),
                    trControl = trControl,
                    metric = "Accuracy",
                    tuneGrid =
                        expand.grid(C = seq(0.01, 3, length = 15),
                                    scale = seq(0.001, 1, length = 15),
                                    degree = as.integer(seq(1, 3, 1))),
                    importance = 'impurity')

# Random Forest
best_tune <- data.frame(
  mtry = numeric(0),
  Num_Trees = numeric(0),
  Split_Rule = character(0),
  Precision = numeric(0),
  Node.Size = numeric(0))

my_seq <- seq(350, 700, 25)
set.seed(123)
for (x in my_seq){
  RF_Model <- train(frmla,
                      PCA_Coord,
                      method = "ranger",
                      trControl = trControl,
                      tuneGrid =
                        expand.grid(.mtry = seq(1, 10, 1),
                                    .min.node.size = seq(1, 6, 1),
                                    .splitrule = c("gini", "extratrees")),
                      metric = "Accuracy",
                      importance = 'impurity')

  Bst_R <- data.frame(
    mtry = RF_Model$bestTune[[1]],
    Num_Trees = x,
    Split_Rule = RF_Model$bestTune[[2]],
    Precision = max(RF_Model$results[[4]]),
    Node.Size = RF_Model$bestTune[[3]]
  )

  best_tune <- rbind(best_tune, Bst_R)

  Bst_R <- c()
}

# Best tune
# mtry = 7; 550 trees split_Rule = extratrees; min.nod.size = 6

```

```

set.seed(123)
RF_Model <- train(
  frmla,
  PCA_Coord,
  method = "ranger",
  trControl = trControl,
  tuneGrid = expand.grid(
    .mtry = 7,
    .min.node.size = 6,
    .splitrule = "extratrees"),
  num.trees = 550,
  metric = "Accuracy",
  importance = 'impurity')

# Boosted tree
set.seed(123)
Boost_Tree <- train(frmla,
                      PCA_Coord,
                      method = "gbm",
                      trControl = trControl,
                      metric = "Accuracy",
                      tuneGrid =
                        expand.grid(
                          n.trees = seq(from = 300, to = 700, by = 50),
                          interaction.depth = seq(from = 1, to = 10, length.out = 5),
                          shrinkage = 0.1,
                          n.minobsinnode = as.integer(seq(1, 10, length = 5)))))

# Multilayer ANN
set.seed(123)
mlp_Mod = train(frmla,
                 PCA_Coord,
                 method = "mlpML",
                 preProc = c('center', 'scale'),
                 trControl = trControl,
                 tuneGrid =
                   expand.grid(
                     layer1 = c(1:8),
                     layer2 = c(0:8),
                     layer3 = c(0:8)))

# Naive Bayes
set.seed(123)
NaiB_Model <- train(frmla,
                      PCA_Coord,
                      method = "nb",
                      preProcess = c("scale", "center"),
                      trControl = trControl,
                      metric = "Accuracy",
                      lineout = FALSE)

confusionMatrix(NaiB_Model)

```

```

# C5.0 Tree
grid <- expand.grid(
  winnow = c(TRUE),
  trials = seq(10, 40, by = 5),
  model = "tree")

set.seed(123)
C50_Mod <- train(frmla,
  PCA_Coord,
  method = "C5.0",
  trControl = trControl,
  metric = "Accuracy",
  importance = 'impurity')

```

## 3 Results

### 3.1 PCA and model performance

The PCA results show that the first 25 PCs account for 95% of the variance of the dataset, with PC1 accounting for 21.39% of the variance and PC25 accounting for 0.36% of the variance. This is an important reduction from the original number of variables (1,524) and is substantially lower than the sample (139).

```

# PCA plot

# Place into data frmae
PC.Plot <- summary(pca)$importance[2:3, 1:25] %>%
  as.data.frame() %>% t() %>%
  as.data.frame()

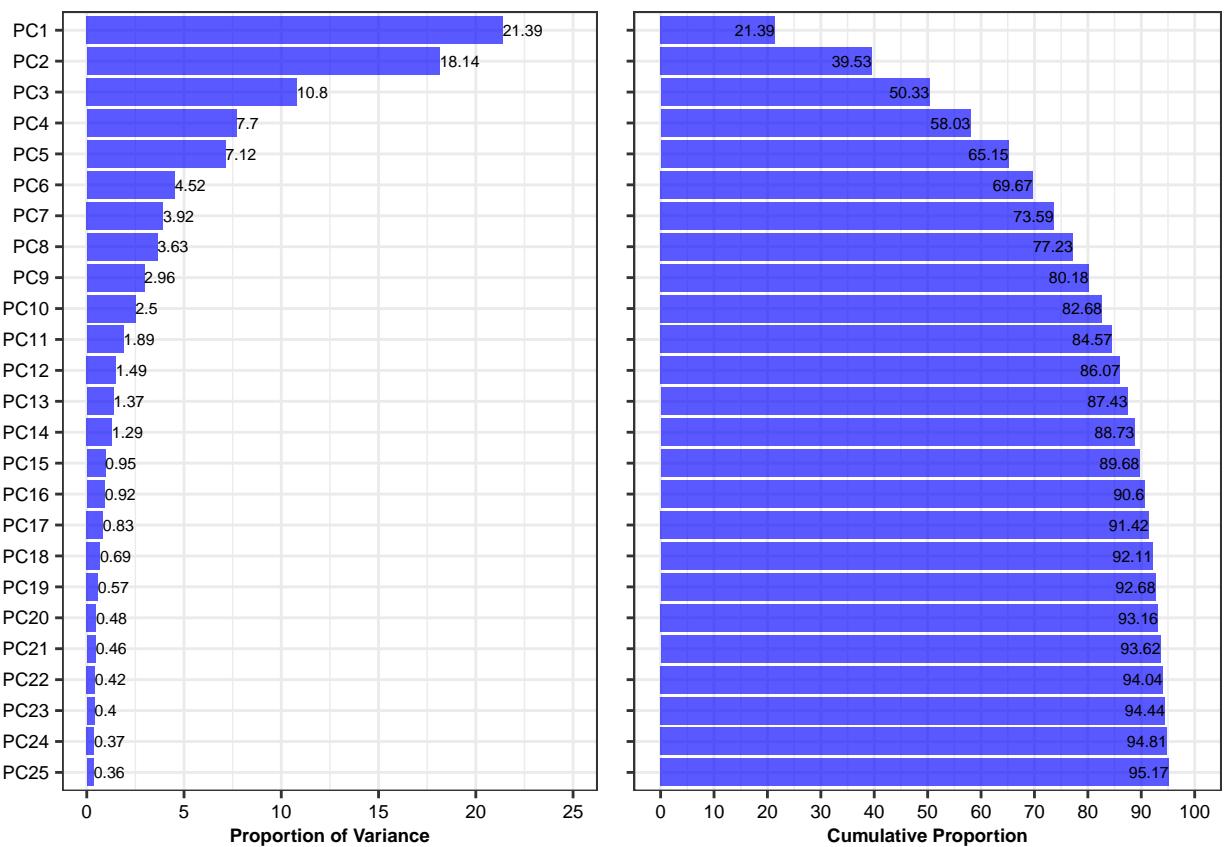
# Bar plot of variance and cumulative variance
ggpubr::ggarrange(
  (PC.Plot %>%
    mutate(PC = rownames(PC.Plot),
      `Proportion of Variance` = `Proportion of Variance`*100) %>%
    ggplot(aes(`Proportion of Variance`, reorder(PC, `Proportion of Variance`))) +
    geom_col(fill = "blue", alpha = 0.65) +
    geom_text(aes(label = round(`Proportion of Variance`, 2)), hjust = "bottom", size = 2.25) +
    theme_bw() +
    scale_x_continuous(breaks = seq(0, 25, 5), lim = c(0, 25)) +
    ylab(NULL) +
    theme(
      axis.title = element_text(size = 7, color = "black", face = "bold"),
      axis.text = element_text(size = 7, color = "black"))
  ),
  (PC.Plot %>%
    mutate(PC = rownames(PC.Plot),
      `Cumulative Proportion` = `Cumulative Proportion`*100) %>%
    ggplot(aes(`Cumulative Proportion`, reorder(PC, -`Cumulative Proportion`))) +

```

```

geom_col(fill = "blue", alpha = 0.65) +
geom_text(aes(label = round(`Cumulative Proportion`, 2)), hjust = "top", size = 2.25) +
theme_bw() +
scale_x_continuous(breaks = seq(0, 100, 10), lim = c(0, 100)) +
ylab(NULL) +
theme(
  axis.title = element_text(size = 7, color = "black", face = "bold"),
  axis.text.x = element_text(size = 7, color = "black"),
  axis.text.y = element_blank())
),
ncol = 2)

```



The following figure presents the performance metrics for each of the models. In general, all models performed with accuracy values higher than 0.7 with the exception of KNN, Naïve Bayes, and the decision tree with C5.0 algorithm. When considering the two measures of overall model performance (F1 and accuracy), SVM with polynomial kernel presents the highest performance values (F1 = 0.75 and accuracy = 0.757). Additionally, SVM with polynomial kernel also provides the highest values of precision.

```

# Data frame of models performance
Model.Performance <- data.frame(
  Model = c("LDA", "KNN", "Log. Reg.", "SVML", "SVMP", "SVMR",
            "C5.0", "Rand. Forest", "GBM", "Naïve Bayes",
            "ANN") %>%
  cbind(
  data.frame(
  rbind(

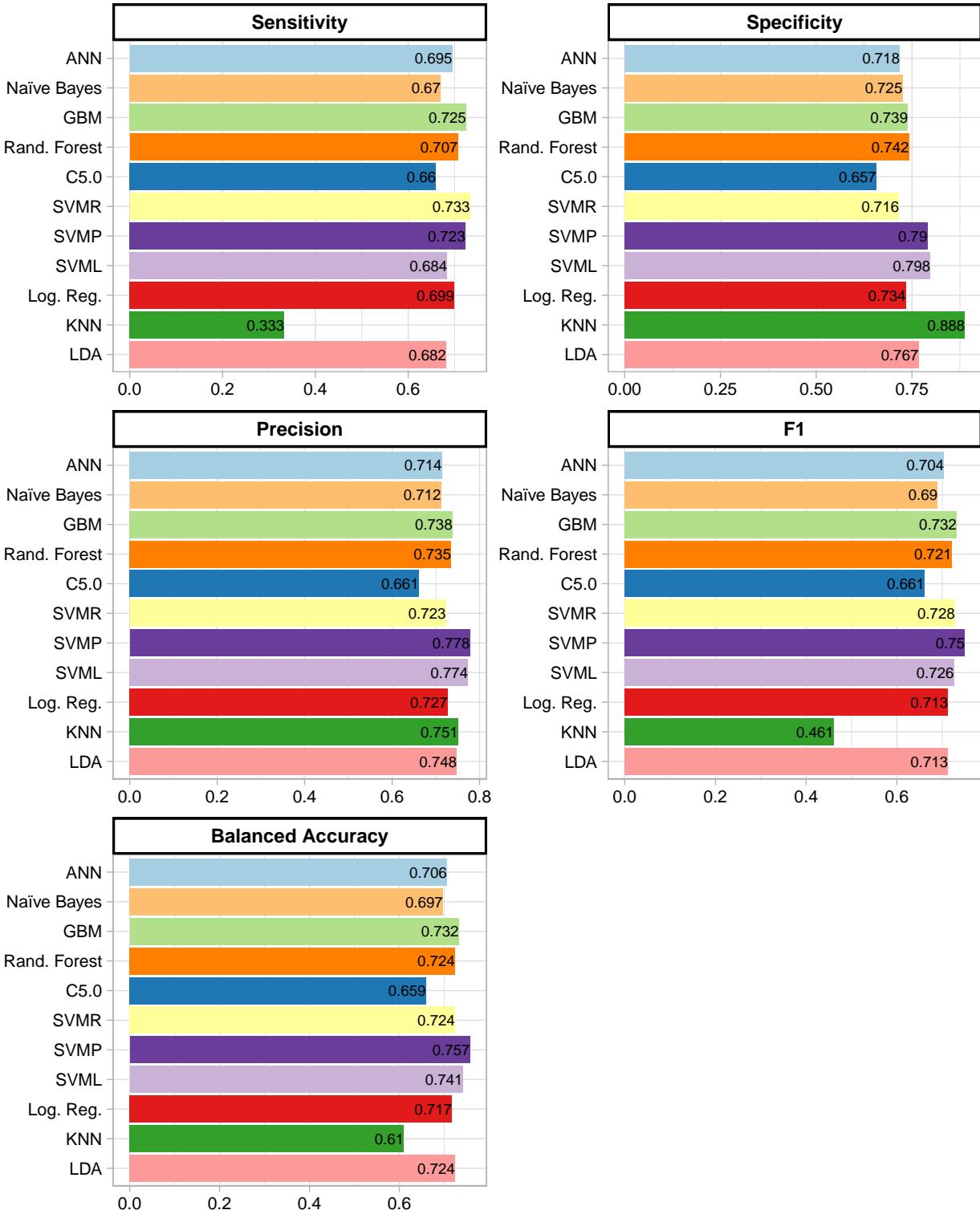
```

```

round(confusionMatrix(fit.LDA$pred$pred, fit.LDA$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(KNN.model$pred$pred, KNN.model$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(logmod$pred$pred, logmod$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(SVM_Linear$pred$pred, SVM_Linear$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(SVM_Poly$pred$pred, SVM_Poly$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(SVM_Radial$pred$pred, SVM_Radial$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(C50_Mod$pred$pred, C50_Mod$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(RF_Model$pred$pred, RF_Model$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(Boost_Tree$pred$pred, Boost_Tree$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(NaiB_Model$pred$pred, NaiB_Model$pred$obs)[[4]][c(1,2,5,7,11)],3),
round(confusionMatrix(mlp_Mod$pred$pred, mlp_Mod$pred$obs)[[4]][c(1,2,5,7,11)],3)))
))

# Plot model performance per metric
Model.Performance %>%
  pivot_longer(cols = Model.Sensitivity:Model.Balanced.Accuracy,
               names_to = "Metric",
               values_to = "Values") %>%
  ggplot(aes(Values, Model.., fill = Model..)) +
  geom_col() +
  facet_wrap(~ factor(Metric,
                       levels = c("Model.Sensitivity",
                                  "Model.Specificity",
                                  "Model.Precision",
                                  "Model.F1",
                                  "Model.Balanced.Accuracy"),
                       labels = c(
                         "Sensitivity",
                         "Specificity",
                         "Precision",
                         "F1",
                         "Balanced Accuracy")),
             scales = "free", ncol = 2) +
  geom_text(aes(label = (Values)), hjust= "top", size = 2.5) +
  theme_light() +
  scale_y_discrete(
    limits = c("LDA", "KNN", "Log. Reg.", "SVML", "SVMP", "SVMR",
              "C5.0", "Rand. Forest", "GBM", "Naïve Bayes",
              "ANN")) +
  scale_fill_brewer(palette = "Paired") +
  ylab(NULL) +
  xlab(NULL) +
  theme(
    strip.text = element_text(color = "black", face = "bold", size = 9),
    strip.background = element_rect(fill = "white", colour = "black", size = 1),
    axis.text = element_text(size = 8, color = "black"),
    legend.position = "none"
  )
)

```



SVM with polynomial kernel is closely followed by SVM with a linear kernel, which presents the second highest value of accuracy (0.741), the fourth highest value of F1 (0.726), and the second-highest value of precision (0.774). Outside SVM with different kernels, the boosted trees also presents high values of accuracy (0.732), F1 (0.732), and precision (0.738). KNN presented the lowest values on the general performance metrics, with an accuracy value of 0.61 and a very low F1 score (0.461). KNN does seem to present high

values of precision (0.751) and specificity (0.888) although these are clearly the result of a sensitivity (0.333) lower than the no-information ratio (0.504).

```

# LDA
temp <- pROC::roc(fit.LDA$pred$obs, fit.LDA$pred$Levallois, quiet = TRUE)
Roc_Curve <- tibble(temp$specificities, temp$sensitivities)
Roc_Curve$Model <- "LDA"

# KNN
temp <- pROC::roc(KNN.model$pred$obs, KNN.model$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "KNN")
Roc_Curve <- rbind(Roc_Curve, temp)

# Log
temp <- pROC::roc(logmod$pred$obs, logmod$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "Log. Reg.")
Roc_Curve <- rbind(Roc_Curve, temp)

# SVML
temp <- pROC::roc(SVM_Linear$pred$obs, SVM_Linear$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "SVML")
Roc_Curve <- rbind(Roc_Curve, temp)

# SVMP
temp <- pROC::roc(SVM_Poly$pred$obs, SVM_Poly$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "SVMP")
Roc_Curve <- rbind(Roc_Curve, temp)

# SVMR
temp <- pROC::roc(SVM_Radial$pred$obs, SVM_Radial$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "SVMR")
Roc_Curve <- rbind(Roc_Curve, temp)

# C5.0
temp <- pROC::roc(C50_Mod$pred$obs, C50_Mod$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "C5.0")
Roc_Curve <- rbind(Roc_Curve, temp)

# Random Forests
temp <- pROC::roc(RF_Model$pred$obs, RF_Model$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "Rand. Forest")
Roc_Curve <- rbind(Roc_Curve, temp)

# GBM
temp <- pROC::roc(Boost_Tree$pred$obs, Boost_Tree$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "GBM")

```

```

Roc_Curve <- rbind(Roc_Curve, temp)

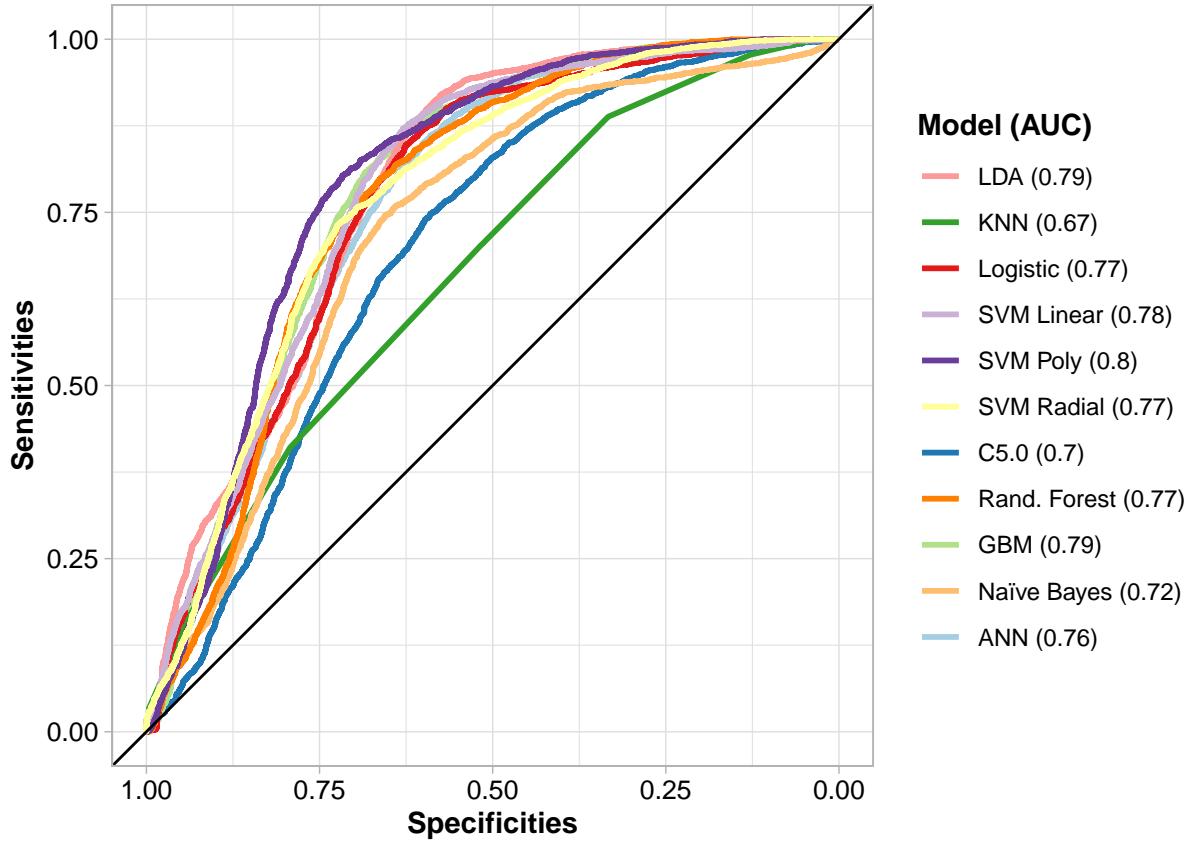
# Boosted tree
temp <- pROC::roc(NaiB_Model$pred$obs, NaiB_Model$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "Naïve Bayes")
Roc_Curve <- rbind(Roc_Curve, temp)

# Boosted tree
temp <- pROC::roc(mlp_Mod$pred$obs, mlp_Mod$pred$Levallois, quiet = TRUE)
temp <- cbind(tibble(temp$specificities, temp$sensitivities),
              Model = "ANN")
Roc_Curve <- rbind(Roc_Curve, temp)
rm(temp)

aucs <- c(
  paste0("LDA (", round(pROC::auc(fit.LDA$pred$obs, fit.LDA$pred$Levallois), 2) ,")"),
  paste0("KNN (", round(pROC::auc(KNN.model$pred$obs, KNN.model$pred$Levallois), 2) ,")"),
  paste0("Logistic (", round(pROC::auc(logmod$pred$obs, logmod$pred$Levallois), 2) ,")"),
  paste0("SVM Linear (", round(pROC::auc(SVM_Linear$pred$obs, SVM_Linear$pred$Levallois), 2) ,")"),
  paste0("SVM Poly (", round(pROC::auc(SVM_Poly$pred$obs, SVM_Poly$pred$Levallois), 2) ,")"),
  paste0("SVM Radial (", round(pROC::auc(SVM_Radial$pred$obs, SVM_Radial$pred$Levallois), 2) ,")"),
  paste0("C5.0 (", round(pROC::auc(C50_Mod$pred$obs, C50_Mod$pred$Levallois), 2) ,")"),
  paste0("Rand. Forest (", round(pROC::auc(RF_Model$pred$obs, RF_Model$pred$Levallois), 2) ,")"),
  paste0("GBM (", round(pROC::auc(Boost_Tree$pred$obs, Boost_Tree$pred$Levallois), 2) ,")"),
  paste0("Naïve Bayes (", round(pROC::auc(NaiB_Model$pred$obs, NaiB_Model$pred$Levallois), 2) ,")"),
  paste0("ANN (", round(pROC::auc(mlp_Mod$pred$obs, mlp_Mod$pred$Levallois), 2) ,")"))

Roc_Curve %>%
  ggplot(aes(`temp$specificities`, `temp$sensitivities`,
             color = Model), alpha = 0.7) +
  geom_line(size = 1.01) +
  scale_x_continuous(trans = "reverse") +
  coord_fixed() +
  theme_light() +
  xlab("Specificities") +
  ylab("Sensitivities") +
  geom_abline(intercept = 1, slope = 1) +
  scale_color_brewer(palette = "Paired",
                     breaks = c("LDA", "KNN", "Log. Reg.",
                               "SVML", "SVMP", "SVMR", "C5.0",
                               "Rand. Forest", "GBM",
                               "Naïve Bayes", "ANN"),
                     labels = aucs) +
  labs(colour = "Model (AUC)") +
  theme(
    axis.title = element_text(size = 11, color = "black", face = "bold"),
    axis.text = element_text(size = 10, color = "black"),
    legend.title = element_text(face = "bold"))

```



The evaluation of the models through the ROC curve and AUC shows that most models present acceptable/fair (0.8–0.7) values. Again, KNN presents the lowest AUC (0.67), a poor value. SVM with polynomial kernel presents the highest AUC value (0.799) and is thus very close to being an excellent/good model (0.9 to 0.8). The optimal probability threshold values from the SVM with polynomial kernel are 0.501 for discoidal and 0.491 for Levallois. The general performance metrics (F1 and accuracy) and AUC values indicate that SVM with polynomial kernel is the best model. The evaluation of SVM with the polynomial kernel confusion matrix shows a very good distribution along the diagonal axis, with the correct identification of Levallois products being slightly higher than the correct identification of discoidal products. The directionality of confusions shows that for the SVM with polynomial kernel, it is more common to mistake discoidal backed products for Levallois rather than mistaking Levallois backed products for those from discoidal reduction sequences.

```
# Confusion matrix
SVM_Poly.Confx <- confusionMatrix(SVM_Poly)$table
SVM_Poly.Confx <- reshape2::melt(SVM_Poly.Confx)

# Normalize the confusion matrix
SVM_Poly.Confx <- SVM_Poly.Confx %>% mutate(
  value = case_when(
    Reference == "Discoid" ~ (value/sum(confusionMatrix(SVM_Poly)$table[1:2]))*100,
    Reference == "Levallois" ~ (value/sum(confusionMatrix(SVM_Poly)$table[3:4]))*100
  )
)

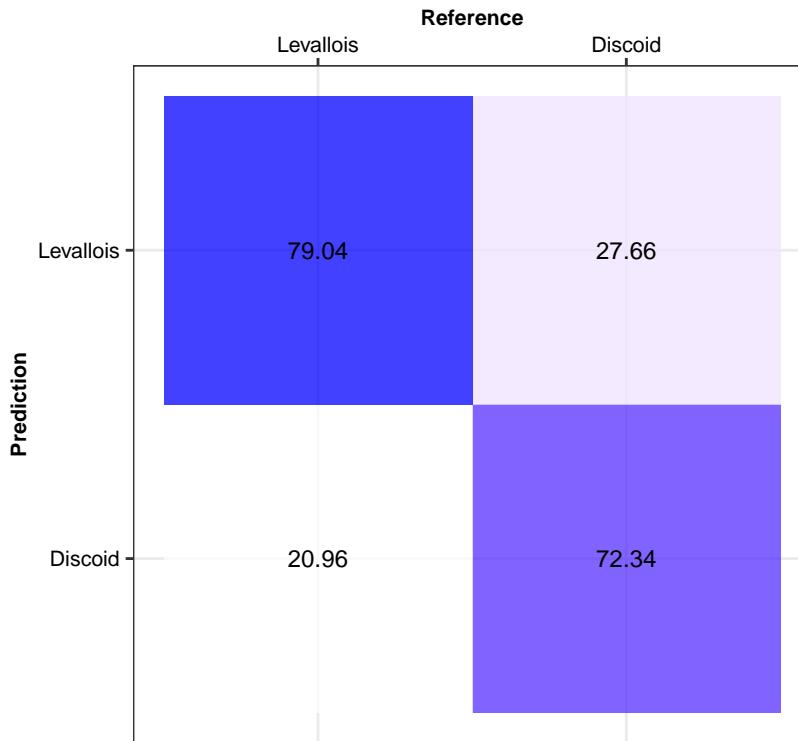
# Set factors
SVM_Poly.Confx$Prediction <- factor(SVM_Poly.Confx$Prediction,
```

```

        levels = c(
            "Discoid", "Levallois"))
SVM_Poly.Confx$Reference <- factor(SVM_Poly.Confx$Reference,
                                     levels = c(
                                         "Levallois", "Discoid"))

# Plot Confusion matrix
SVM_Poly.Confx %>%
  ggplot(aes(Reference, Prediction, fill = value)) +
  geom_tile(alpha = 0.75) +
  geom_text(aes(label = round(value, 2)), size = 3) +
  scale_fill_gradient(low = "white", high = "blue") +
  scale_x_discrete(position = "top") +
  theme_bw() +
  coord_fixed() +
  theme(legend.position = "none",
        axis.title = element_text(size = 8, color = "black", face = "bold"),
        axis.text = element_text(size = 7.5, color = "black"),
        title = element_text(size = 8, color = "black", face = "bold"))

```



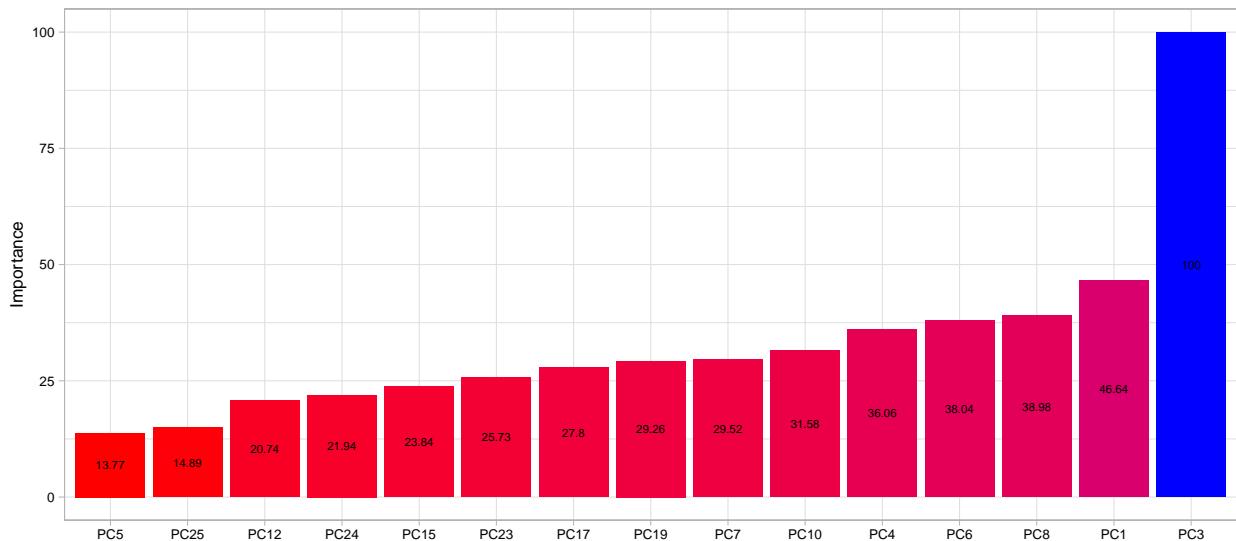
### 3.2 PC Importance

The following figure presents the PC importance for the discrimination of knapping method according to SVM with polynomial kernel model. The PC importance shows that PC3 clearly stands out in importance for the discrimination of discoidal and Levallois backed products. PC3 only accounts for 10.8% of the variance but presents the maximum scaled importance. PC1, which represents 21.39% of the variance, is the second most important variable, with a score of 46.64, although far from PC3. PC8, which represents only 3.63% of the variance, is the third most important variable for the SVM with polynomial kernel model.

```

# Data frame of PC importance
tibble(
  PC = rownames(varImp(SVM_Poly, sale = TRUE)$importance),
  Importance = varImp(SVM_Poly, sale = TRUE)$importance[, 1]) %>%
  
  top_n(15, Importance) %>%
# and plot
  ggplot(aes(Importance, reorder(PC, Importance), fill = Importance)) +
  geom_bar(stat = "identity", position = "dodge") +
  geom_text(aes(label = round(Importance, 2)),
            position = position_stack(vjust = 0.5), size = 2) +
  scale_fill_gradient(low = "red", high = "blue") +
  guides(fill = "none") +
  coord_flip() +
  ylab(NULL) +
  theme_light() +
  theme(
    axis.text.y = element_text(color = "black", size = 7),
    axis.text.x = element_text(color = "black", size = 7),
    axis.title.x = element_text(color = "black", size = 9),
    axis.title.y = element_text(color = "black", size = 9))

```



```

#PC biplots
ggpubr::ggarrange(
(PCA_Coord %>% ggplot(aes(PC3, PC1, fill = Strategy)) +
  geom_vline(xintercept = 0, alpha = 0.7, linetype = "dashed") +
  geom_hline(yintercept = 0, alpha = 0.7, linetype = "dashed") +
  stat_ellipse(geom = "polygon", alpha = 0.2, aes(fill = Strategy)) +
  geom_point(aes(color = Strategy)) +
  xlab(paste0("PC3 (", round((summary(pca)$importance[2,3])*100, 2), "%)")) +
  ylab(paste0("PC1 (", round((summary(pca)$importance[2,1])*100, 2), "%)")) +
  ggsci::scale_fill_aaas() +
  ggsci::scale_color_aaas() +
  theme_light() +
  theme(

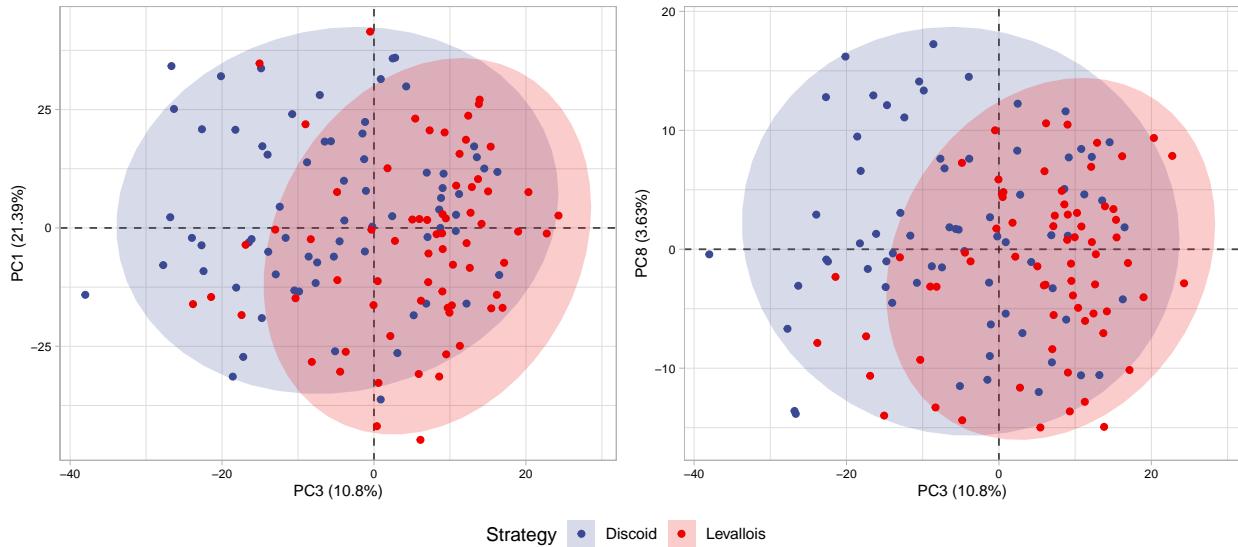
```

```

axis.text.y = element_text(color = "black", size = 7),
axis.text.x = element_text(color = "black", size = 7),
axis.title.x = element_text(color = "black", size = 9),
axis.title.y = element_text(color = "black", size = 9),
legend.position = "bottom"),

(PCA_Coord %>% ggplot(aes(PC3, PC8, fill = Strategy)) +
  geom_vline(xintercept = 0, alpha = 0.7, linetype = "dashed") +
  geom_hline(yintercept = 0, alpha = 0.7, linetype = "dashed") +
  stat_ellipse(geom = "polygon", alpha = 0.2, aes(fill = Strategy)) +
  geom_point(aes(color = Strategy)) +
  xlab(paste0("PC3 (", round((summary(pca)$importance[2,3])*100, 2), "%)")) +
  ylab(paste0("PC8 (", round((summary(pca)$importance[2,8])*100, 2), "%)")) +
  ggsci::scale_fill_aaas() +
  ggsci::scale_color_aaas() +
  theme_light() +
  theme(
    axis.text.y = element_text(color = "black", size = 7),
    axis.text.x = element_text(color = "black", size = 7),
    axis.title.x = element_text(color = "black", size = 9),
    axis.title.y = element_text(color = "black", size = 9),
    legend.position = "bottom")),
  ncol = 2,
  common.legend = TRUE,
  legend = "bottom",
  align = "h")

```

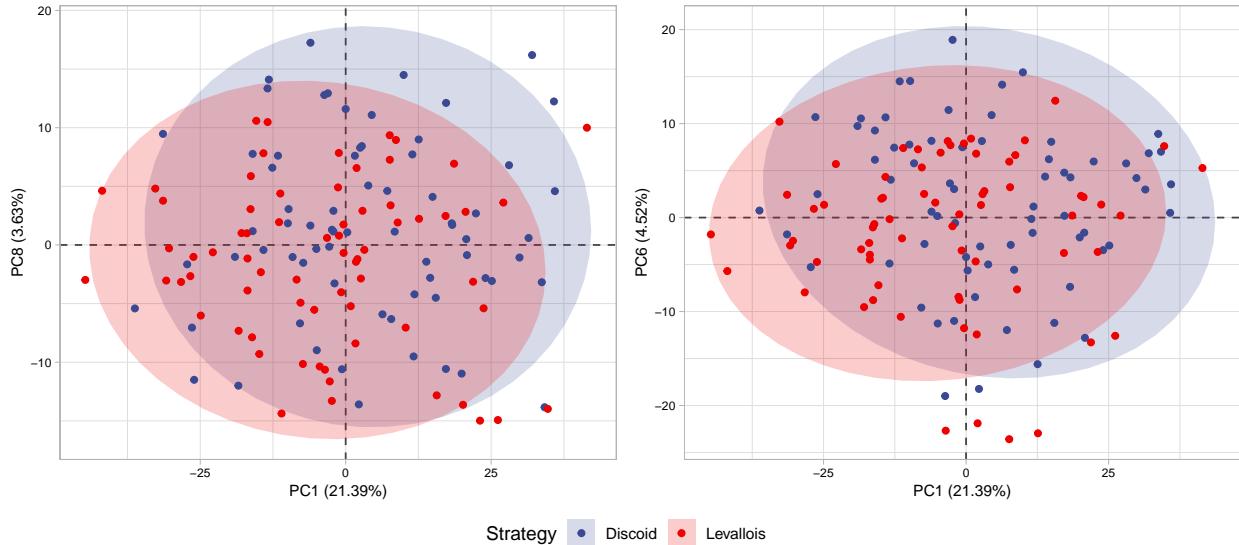


The effect of PC3 on identifying backed products from the two knapping methods is especially notable in the a biplot distribution. The above figure presents a biplot distribution of the data between PC3 and the following two most important variables. In both cases, backed flakes detached from Levallois recurrent centripetal cores tend to be clustered in the positive values of PC3, whereas they show a wider distribution, usually centered on the 0 value, for PCs 1 and 8. Backed flakes from discoidal reduction sequences show a wider distribution although the center is in the negative and 0 values of PC3. Although the combination of PC3 with PC1 and PC8 shows an overlapping of the confidence ellipsis, differentiation between both groups

can be observed.

The following figure presents a biplot distribution of the data when the second (PC1), third (PC8), and fourth (PC6) most important variables are employed. The biplot from the combination of these variables shows much more consistent overlapping for the different combinations of PC1, PC8, and PC6.

```
ggpubr::ggarrange(
  (PCA_Coord %>% ggplot(aes(PC1, PC8, fill = Strategy)) +
    geom_vline(xintercept = 0, alpha = 0.7, linetype = "dashed") +
    geom_hline(yintercept = 0, alpha = 0.7, linetype = "dashed") +
    stat_ellipse(geom = "polygon", alpha = 0.2, aes(fill = Strategy)) +
    geom_point(aes(color = Strategy)) +
    xlab(paste0("PC1 (", round((summary(pca)$importance[2,1])*100, 2), "%)")) +
    ylab(paste0("PC8 (", round((summary(pca)$importance[2,8])*100, 2), "%)")) +
    ggsci::scale_fill_aaas() +
    ggsci::scale_color_aaas() +
    theme_light() +
    theme(
      axis.text.y = element_text(color = "black", size = 7),
      axis.text.x = element_text(color = "black", size = 7),
      axis.title.x = element_text(color = "black", size = 9),
      axis.title.y = element_text(color = "black", size = 9),
      legend.position = "bottom")),
  (PCA_Coord %>% ggplot(aes(PC1, PC6, fill = Strategy)) +
    geom_vline(xintercept = 0, alpha = 0.7, linetype = "dashed") +
    geom_hline(yintercept = 0, alpha = 0.7, linetype = "dashed") +
    stat_ellipse(geom = "polygon", alpha = 0.2, aes(fill = Strategy)) +
    geom_point(aes(color = Strategy)) +
    xlab(paste0("PC1 (", round((summary(pca)$importance[2,1])*100, 2), "%)")) +
    ylab(paste0("PC6 (", round((summary(pca)$importance[2,6])*100, 2), "%)")) +
    ggsci::scale_fill_aaas() +
    ggsci::scale_color_aaas() +
    theme_light() +
    theme(
      axis.text.y = element_text(color = "black", size = 7),
      axis.text.x = element_text(color = "black", size = 7),
      axis.title.x = element_text(color = "black", size = 9),
      axis.title.y = element_text(color = "black", size = 9),
      legend.position = "bottom")),
  ncol = 2,
  common.legend = TRUE,
  legend = "bottom",
  align = "h")
```



Prior to perform multiple linear regression to predict PC values based on flake metric features it is necessary to join both data sets (the one containing PC values and the one containing attribute analysis values). The following code load the data from recorded attributes and performs a `left_join()` to match PC and attribute values according to flake ID.

```
# Read in attribute dataset
Att <- read.csv("Data/Attributes data.csv")

# Left joined with the attribute database
PCA_Coord <- left_join(PCA_Coord, Att, by = "ID")

# Compute ratios
PCA_Coord <- PCA_Coord %>%
  mutate(Lam.Ind = LENGTH/WIDTH,
        Caren.Ind = case_when(
          LENGTH < WIDTH ~ LENGTH/MAXTHICK,
          LENGTH > WIDTH ~ WIDTH/MAXTHICK,
          LENGTH == WIDTH ~ WIDTH/MAXTHICK),
        Flat_Ind = (WIDTH*LENGTH)/MAXTHICK,
        Flak.Surface = WIDTH*LENGTH,
        W.to.Thic = WIDTH/MAXTHICK)
```

Multiple linear regression for the prediction of PC3 indicates that the best correlation is obtained when the interaction of IPA and the ratio of flake width to thickness is employed ( $p < 0.001$ , adjusted  $r^2 = 0.65$ ). The coefficient of the interaction between the ratio of width to thickness and IPA is 0.17, whereas the coefficient of IPA is -0.77. This indicates that as the IPA becomes more open as the values of PC3 decrease. The ratio of flake width to thickness offers a counterintuitive coefficient of -12.79. The signal of this coefficient is opposite to that obtained from a linear regression where the values of the ratio of flake width to thickness are employed to predict PC3 values ( $p < 0.001$ ;  $r^2 = 0.6$ ; coefficient = 6.46). The reversed signal obtained from the interaction can be considered the result of Simpson's paradox (Simpson, 1951). The high correlation between the carenated index and the ratio of flake width to thickness ( $p < 0.001$ ;  $r^2 = 0.9$ ) indicates that PC3 captures relative flake thinness to thickness although it regresses better with the ratio of width to thickness. In general, thin flakes with an IPA close to 90° will have high positive PC3 values, whereas thick flakes with open IPA will have negative values.

```

# Best predictors for PC3
summary(lm(PC3 ~ W.to.Thic*IPA, PCA_Coord))

## 
## Call:
## lm(formula = PC3 ~ W.to.Thic * IPA, data = PCA_Coord)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -22.4659  -4.6226  -0.2159   4.8543  21.1260 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 62.02025  17.75208  3.494 0.000645 ***
## W.to.Thic   -12.79366  4.22008 -3.032 0.002917 ** 
## IPA        -0.77034  0.15657 -4.920 2.48e-06 *** 
## W.to.Thic:IPA  0.17362  0.03855  4.503 1.43e-05 *** 
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 7.597 on 135 degrees of freedom
## Multiple R-squared:  0.657, Adjusted R-squared:  0.6494 
## F-statistic: 86.21 on 3 and 135 DF, p-value: < 2.2e-16

```

```

# Correlation between Carenated index and ratio of width to thickness
summary(lm(Caren.Ind ~ W.to.Thic, PCA_Coord))

```

```

## 
## Call:
## lm(formula = Caren.Ind ~ W.to.Thic, data = PCA_Coord)
## 
## Residuals:
##      Min       1Q   Median       3Q      Max 
## -1.74506 -0.15789  0.09566  0.26291  0.86470 
## 
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)    
## (Intercept) 0.21885  0.10253  2.134   0.0346 *  
## W.to.Thic   0.86108  0.02486 34.632  <2e-16 *** 
## ---      
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 
## 
## Residual standard error: 0.4477 on 137 degrees of freedom
## Multiple R-squared:  0.8975, Adjusted R-squared:  0.8967 
## F-statistic: 1199 on 1 and 137 DF, p-value: < 2.2e-16

```

An analysis of PC3 values according to group shows that backed products from Levallois recurrent centripetal methods tend to have higher values (mean = 5.47) with a slightly lower standard deviation. Backed products detached from discoidal cores, alternatively, tend to have lower values (mean = -5.20) and a slightly higher standard deviation (12.74).

```

# Descriptive statistics of PC3
data.frame(
  PCA_Coord %>%
    group_by(Strategy.y) %>%
    summarise(
      Min = min(PC3),
      `5th Perc` = quantile(PC3, 0.05),
      `1st quantile` = quantile(PC3, 0.25),
      Mean = mean(PC3),
      Median = quantile(PC3, 0.5),
      `3rd quantile` = quantile(PC3, 0.75),
      `95 Perc` = quantile(PC3, 0.95),
      Max = max(PC3),
      SD = sd(PC3)))

```

```

##   Strategy.y      Min X5th.Perc X1st.quantile      Mean     Median X3rd.quantile X95.Perc
## 1   Discoid -38.00047 -26.50603    -14.741836 -5.388529 -5.200406 4.984441 13.3835
## 2  Levallois -23.82052 -16.16213      0.385224  5.466624  8.588994 12.411823 18.2428
##       Max        SD
## 1 16.49057 12.74186
## 2 24.29432 10.44742

```

Multiple linear regression for the prediction of PC1 values shows a moderate correlation when the elongation index and carenated index are employed as predictors ( $p < 0.001$ ; adjusted  $r^2 = 0.63$ ). The elongation index presents the highest significance and the highest estimate value (-39.27), whereas the carenated index presents an estimate value of -4.26. The negative and high value of the estimate for the elongation index indicates that as the elongation tendency of a product increases (becoming longer relative to its width), the values of PC1 will decrease while all other variables remain constant. The negative estimate of the carenated index also indicates that as a product becomes thinner, the values of PC1 will decrease. Thus, the positive values of PC1 represent thick products with a low elongation.

```

# Prediction of PC1 values
summary(lm(PC1 ~ Caren.Ind + Lam.Ind, PCA_Coord))

##
## Call:
## lm(formula = PC1 ~ Caren.Ind + Lam.Ind, data = PCA_Coord)
##
## Residuals:
##      Min      1Q      Median      3Q      Max
## -34.306  -5.736   0.297   8.171  22.534
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept) 59.0190    4.3058 13.707 < 2e-16 ***
## Caren.Ind   -4.2556    0.6908 -6.161 7.69e-09 ***
## Lam.Ind     -39.2654   2.6099 -15.045 < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 11.02 on 136 degrees of freedom
## Multiple R-squared:  0.633, Adjusted R-squared:  0.6276
## F-statistic: 117.3 on 2 and 136 DF, p-value: < 2.2e-16

```

The analysis of PC1 values shows differences between the backed products of the discoidal and Levallois recurrent centripetal methods. On average, backed products from the Levallois recurrent centripetal method will have higher values (mean = 3.75) compared to discoidal products (mean = -3.80). However, an important overlapping of values is evident for products from both reduction methods, with high values of standard deviation in both cases.

```
# Descriptive statistics of PC1
data.frame(
  PCA_Coord %>%
    group_by(Strategy.y) %>%
    summarise(
      Min = min(PC1),
      `5th Perc` = quantile(PC1, 0.05),
      `1st quantile` = quantile(PC1, 0.25),
      Mean = mean(PC1),
      Median = quantile(PC1, 0.5),
      `3rd quantile` = quantile(PC1, 0.75),
      `95 Perc` = quantile(PC1, 0.95),
      Max = max(PC1),
      SD = sd(PC1)))
##   Strategy.y      Min X5th.Perc X1st.quantile      Mean     Median X3rd.quantile X95.Perc
## 1   Discoid -36.23377 -26.25343     -7.711467  3.747729  2.372406    16.789563 32.95855
## 2  Levallois -44.75778 -31.17264    -16.301620 -3.802044 -2.736412     7.584384 25.18118
##           Max      SD
## 1 35.95467 17.3134
## 2 41.46336 18.1138
```

## 4. Discussion

Our results have shown an accuracy of 0.76 for the differentiation of backed pieces from discoidal and Levallois recurrent centripetal methods. Additionally, the use of decision thresholds provided an AUC close to 0.8. This degree of accuracy indicates that the quantification of morphological features through geometric morphometrics, along with dimensionality reduction using PCA and machine learning models, can accurately differentiate between the two methods tested. Of the 11 models tested, SVM with polynomial kernel provided the best performance for the discrimination of discoidal and Levallois recurrent centripetal methods in backed artifacts. Moreover, results support the notion that discoidal and recurrent centripetal Levallois are two separate core reduction methods/conceptions.

The selection of machine learning algorithms depends on the task, type of classification, and nature of the data, requiring for the systematical evaluation of different algorithms. Of the tested models, SVM with polynomial kernel performed the best. SVM's have a series of features which make them ideal for the analysis of lithic artifacts. As previously mentioned, SVM models fit hyperplanes to separate classes, use margins to find the best separation, and apply a cost value with each misclassification (Lantz, 2019). Selection of the kernel is key for the performance of SVM, since it will have a direct impact on the fitted hyperplane, margins and cost value. As a result of these features, SVM are able to allow misclassifications and overlapping classifications in order to obtain a better general performance. These features of SVM make them especially adequate for the analysis of lithic materials and might be one of the underlying reasons why it has the best performance metrics.

The first 25 PCs captured 95% of the sample variance. Of these PCs, the highest importance value for the discrimination of knapping methods was obtained by PC3. Multiple linear regression shows that PC3 is moderately correlated with an interaction between IPA and the ratio of flake width to thickness. Thin and wide artifacts with IPA values close to 90° will have higher PC3 values. The examination of biplots

and PC3 values shows that backed flakes detached from Levallois recurrent centripetal cores tend to be thin in relation to the thickness, non-elongated, and have an IPA close to 90°. PC1, which captures elongation tendencies with higher resolution (along with product thickness) also supports this interpretation although higher overlapping exists. The discriminatory power of PC3 appears inherently related to differences in how the volume of cores in the two methods are conceived: non-hierarchized surfaces exploited with secant removals in discoidal reduction while recurrent centripetal Levallois is characterized by sub-parallel removals from a single debitage surface. Additional features for the discrimination of discoidal and Levallois backed products can be found in edge angles and the angles of negatives of the dorsal face towards the detachment surface. In general, it is expected that products detached from Levallois reduction sequences will have more acute edge angles, along with flatter dorsal surface negatives. Again, these differences in the angles are also inherently related to differences in how the volume of cores in the two methods are conceived.

These results indicate that there are underlying morphological differences between backed artifacts detached from both methods. These underlying morphological differences can be captured and quantified by geometric morphometrics along with PCA and used by machine learning models for an accurate discrimination of methods.

Several authors have identified underlying morphological characteristics that can differentiate backed products detached from Levallois recurrent centripetal and discoidal cores ([Boëda et al., 1990](#); [Delpiano et al., 2021](#); [Meignen, 1996](#); [Meignen, 1993](#); [Mourre, 2003](#)). As previously pointed out, one of the features captured by PC3 is the IPA. Products that have an IPA close to 90° have increasing PC3 values, being the case for most Levallois backed flakes. Levallois products having an IPA close to 90° have been described previous to ([Kelly, 1954](#)) and after ([Boëda, 1995b](#), [1994](#), [1993](#)) the technological description of the Levallois flaking system. A recent study employing machine learning models ([González-Molina et al., 2020](#)) on attribute analysis have also pointed to the IPA as one of the features allowing differentiation between Levallois recurrent centripetal and Discoid products.

[Delpiano et al. \(2021\)](#) focused on the general morphology of backed products from discoidal and Levallois reduction sequences, stating that the latter tend to be thinner with subparallel and rectilinear edges and a higher elongation index. The interpretation of PC3 in the present study also identifies backed products from Levallois recurrent centripetal as being thinner. However, the interpretation of PC1 (which better captures elongation) proved not to be a sound criterion for discriminating between strategies, with both methods showing a very wide range of elongation values. Concerning the elongation of Levallois products, [Boëda et al. \(1990\)](#) also noticed the decrease of length/width ratio with each successive exploitation, resulting in short non-laminar flakes and core edge flakes. [Mourre \(2003\)](#) equally called attention to the direction of removal axis being parallel to the plane of intersection of both surfaces in the case of Levallois core edge flakes. In our study, the effect of this feature can be linked to a higher carenation index, which is captured by PC3. The visual exploration of the 3D meshes according to PC values did not seem to capture the relation between the debitage axis and the symmetry of blanks as an important feature of Levallois centripetal backed flakes ([Meignen, 1996](#); [Meignen, 1993](#)). This is probably due to the inclusion of core edge flakes with a limited back in the experimental sample and its possible importance being overshadowed by other features better for discrimination captured by PC3 (IPA, carenation index and elongation index).

[González-Molina et al. \(2020\)](#) achieved an 80% accuracy when differentiating between discoidal and Levallois centripetal flakes. Although their study focused uniquely on flakes from the exploitation phase (with the presence of cortex having very little importance as a variable for differentiating methods), it did not specifically addressed the issue of backed flakes, and dimensional variables (width at different points and maximum thickness) have high importance, it shows the potential of using machine learning models for the identification of knapping methods. In contrast, our study focused on a concrete set of technological products independent of the reduction phase, and the use of geometric morphometrics excludes dimensional variables. However, despite these differences, similar degrees of accuracy were obtained. [Archer et al. \(2021\)](#) also used geometric morphometrics and random forest to evaluate the differentiation between three strategies (Levallois, discoidal, and laminar). Although overall performance of the models is based on accuracy, their study nevertheless reached a similar value to that of the present study. However, the classification of the two same classes as in the present study varies significantly, with an 87% accuracy for Levallois products and 40% in the differentiation of discoidal products. This contrasts heavily with our results, where the classification is more balanced and the identification of backed products from discoidal reduction showed a slightly lower

accuracy than the identification of products from the Levallois recurrent centripetal method (0.72 and 0.79, respectively).

Archer et al. (2021) also reported human analyst identification ratios for flakes from different archaeological sites with the “undiagnostic” class being the largest, usually tallying above 60% (thus, only 35% of flakes were attributed to a knapping method). In both the above-mentioned studies (Archer et al., 2021; González-Molina et al., 2020) and in the present study, the application of machine learning models notably increases the accuracy and predictions regarding the identification of knapping methods. Caution is always advisable when evaluating such findings, as controlled experimental assemblages do not mimic the complexity of the archaeological record.

The present study has employed multiple linear regression with common metrics of lithic analysis as predictors to determine what features were captured by the PCs. The multiple linear regressions of both PC3 and PC1 presented moderate correlation values, with more than 0.6 of the variance explained. However, this also implies that a good portion of the variance remains unexplained for both PCs. The remaining unexplained variance can be the result of several factors. Metric variables used as predictors were taken manually, likely resulting in some degree of error. Geometric morphometrics capture the same metric variables with higher resolution, thus representing another potential source of error when establishing correlations. An additional source of the unexplained PC variance might come from metric features (and their interactions) recorded as part of exhaustive attribute analyses (e.g. the number, organization and flaking angle of previous removals). This suggests future research should take into account large sample sizes along with the incorporation of these analytical features.

While backed flakes detached from discoidal and Levallois recurrent centripetal methods were the focus of our analysis, it is important to note that backed products are common to other flaking strategies such as Quina and SSDA (Bourguignon, 1996; Forestier, 1993) not included in the present study. Although in Western Europe the coexistence of Levallois and discoidal knapping methods with other knapping methods in the same archaeological levels is a subject of debate (Faivre et al., 2017; Grimaldi and Santaniello, 2014; Marciani et al., 2020; Ríos-Garaizar, 2017), the present model can be applied to assemblages where Levallois and Discoid knapping strategies have been shown to coexist. For this, the study and evaluation of the chaîne opératoire and assemblage context and integrity are fundamental for the study of lithic technology (Soressi and Geneste, 2011). Thus, the chaîne opératoire and assemblage integrity should be considered prior to the application of geometric morphometrics and machine learning models for the identification of knapping methods.

## 5. Conclusions

Backed flakes are technological products that play special roles in the discoidal and Levallois recurrent centripetal methods (Boëda, 1993; Boëda et al., 1990; Slimak, 2003). In Levallois reductions, these removals serve to manage lateral and distal core convexities (Boëda et al., 1990), while their systematic production in discoidal reductions demonstrates their properties to be intentionally sought-after (Locht, 2003; Slimak, 2003). Additionally, data from several sites show that they were commonly imported and exported (Geneste, 1988; Roebroeks et al., 1992; Turq et al., 2013). This frequent transport is possibly connected to their specific morpho-functional features (Delpiano et al., 2021), a “prehensile” core edge opposite a convergent cutting edge. Associated with two technologically distinct core reduction methods, it should be expected that their morphological features differ and therefore allow for the identification of the knapping method. With the use of geometric morphometrics, these morphological features can be quantified, and PCA for dimensionality reduction allows them to be employed in machine learning models.

PCA and machine learning models indeed capture the different morphological features derived from both knapping methods, resulting in an accuracy of 0.76 and an AUC of 0.8 in the case of the best model for differentiating between knapping strategies. Most of the importance for differentiating between the knapping methods was captured by only one variable (PC3), which multiple linear regressions showed to be correlated with the elongation index and mostly an interaction between IPA and the carenation index.

The workflow of the present work has been developed to allow for its implementation on archaeological materials. Archaeological backed flakes can be scanned, then use the flake template from the present work

to locate landmarks on the 3D meshes, perform procrustes alignment of all backed flakes (archaeological and experimental reference flakes of the present research whose original unaligned coordinates are available), reduce dimensionality of experimental and archaeological backed flakes through PCA (it is important to notice that PCA values will change with the inclusion of new archaeological data), train the SVM with polynomial kernel on the experimental assemblage, and make predictions on the archaeological material. Of this workflow, 3D scanning and landmarks positioning are considered to be the most time-consuming stages, while procrustes alignment, PCA and model training can be considered to be fairly fast.

Geometric morphometrics in combination with dimensionality reduction methods (PCA) and machine learning models can offer high-resolution methods for the identification of knapping methods in lithic analysis although their application should not be independent from the study of the operative chain and assemblage technological context.

## Acknowledgments

This research has been supported by the project SI1/PJI/2019-00488 funded by Comunidad Autónoma de Madrid and Universidad Autónoma de Madrid.

## Author Contributions

FR and GBP conceived and designed the research and the experiments. GBP, BG, MB performed the experiments. GBP analyzed and curated the data. Original draft was written by GBP and FR. MB, FR, and BG reviewed and edited the manuscript. FR was responsible for funding acquisition. All the authors agreed on the final version of the paper.

## References

- Archer, W., Djakovic, I., Brenet, M., Bourguignon, L., Presnyakova, D., Schlager, S., Soressi, M., McPherron, S.P., 2021. Quantifying differences in hominin flaking technologies with 3D shape analysis. *Journal of Human Evolution* 150, 102912. <https://doi.org/10.1016/j.jhevol.2020.102912>
- Archer, W., Pop, C.M., Rezek, Z., Schlager, S., Lin, S.C., Weiss, M., Dogandžić, T., Desta, D., McPherron, S.P., 2018. A geometric morphometric relationship predicts stone flake shape and size variability. *Archaeological and Anthropological Sciences* 10, 1991–2003. <https://doi.org/10.1007/s12520-017-0517-2>
- Bergmeir, C., Benítez, J.M., 2012. Neural Networks in R using the Stuttgart Neural Network Simulator: RSNNS. *Journal of Statistical Software* 46, 1–26. <https://doi.org/10.18637/jss.v046.i07>
- Beyries, S., Boëda, E., 1983. Étude technologique et traces d'utilisation des éclats débordants de Corbehem (Pas-de-Calais). *Bulletin de la Société préhistorique française* 80, 275–279. <https://doi.org/10.3406/bspf.1983.5455>
- Binford, L.R., 1979. Organization and Formation Processes: Looking at Curated Technologies. *Journal of Anthropological Research* 35, 255–273.
- Boëda, E., 1995a. Levallois: A Volumetric Construction, Methods, A Technique, in: Dibble, H.L., Bar-Yosef, O. (Eds.), *The Definition and Interpretation of Levallois Technology*, Monographs in World Archaeology. Prehistory Press, Madison, Wisconsin, pp. 41–68.
- Boëda, E., 1995b. Caractéristiques techniques des chaînes opératoires lithiques des niveaux micoquiens de Külna (Tchécoslovaquie). *Paléo* 1, 57–72. <https://doi.org/10.3406/pal.1995.1380>
- Boëda, E., 1994. Le concept Levallois: Variabilité des méthodes, CNRS éditions. CNRS.
- Boëda, E., 1993. Le débitage discoïde et le débitage Levallois récurrent centripète. *Bulletin de la Société Préhistorique Française* 90, 392–404. <https://doi.org/10.3406/bspf.1993.9669>
- Boëda, E., Geneste, J.-M., Meignen, L., 1990. Identification de chaînes opératoires lithiques du Paléolithique ancien et moyen. *Paléo* 2, 43–80.
- Bookstein, F.L., 1997b. 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](https://doi.org/10.1016/S1361-8415(97)85012-8)

- Bookstein, F.L., 1997a. Morphometric tools for landmark data. Cambridge University Press.
- Bordes, F., 1961. Typologie du Paléolithique ancien et moyen, Publications de l'institut de préhistoire de l'université de Bordeaux. CNRS Editions, Bordeaux.
- Bordes, F., 1953. Notules de typologie paléolithique II : Pointes Levalloisiennes et pointes pseudo-levalloisiennes. Bulletin de la Société préhistorique de France 50, 311–313. <https://doi.org/10.3406/bspf.1953.3057>
- Bourguignon, L., 1996. La conception du debitage Quina. *Quaternaria Nova* 6, 149–166.
- Bradley, A.P., 1997. The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern recognition* 30, 1145–1159.
- Breiman, L., 2001. Random Forests. *Machine Learning* 45, 5–32. <https://doi.org/10.1023/A:1010933404324>
- Brenet, M., 2013. Variabilité et signification des productions lithiques au Paléolithique moyen ancien. L'exemple de trois gisements de plein-air du Bergeracois (Dordogne, France), BAR International Series. Archaeopress, Oxford.
- Brenet, M., 2012. Silex et roches métamorphiques au Paléolithique moyen récent: Combemenu (Corrèze) et Chemin d'Herbe (Lot-et-Garonne), in: Marchand, G., Querré, G. (Eds.), *Roches Et Sociétés de La Préhistoire Entre Massifs Cristallins Et Bassins Sédimentaires : Le Nord-Ouest de La France Dans Son Contexte Européen*. Presses universitaires de Rennes, pp. 379–393.
- Brenet, M., Cretin, C., 2008. Le gisement paléolithique moyen et supérieur de Combemenu (Brignac-la-plaine, Corrèze). Du microvestige au territoire, réflexions sur les perspectives d'une approche multi-scalaire, in: Aubry, T., Almeida, F., Araújo, A.C., Tiffagom, M. (Eds.), *Proceedings of the XV World Congress UISPP* (Lisbon, 4-9 September 2006). Space and Time: Which Diachronies, Which Synchronies, Which Scales? / Typology Vs Technology, Sessions C64 and C65, BAR International Series. Archaeopress, Oxford, pp. 35–44.
- Bustillo, M.A., Pérez-Jiménez, J.L., 2005. Características diferenciales y génesis de los niveles silíceos explotados en el yacimiento arqueológico de Casa Montero (Vicálvaro, Madrid). *Geogaceta* 38, 243–246.
- Bustillo, M.Á., Pérez-Jiménez, J.L., Bustillo, M., 2012. Caracterización geoquímica de rocas sedimentarias formadas por silicificación como fuentes de suministro de utensilios líticos (Mioceno, cuenca de Madrid). *Revista Mexicana de Ciencias Geológicas* 29, 233–247.
- Bustos-Pérez, G., Chacón, M.G., Rivals, F., Blasco, R., Rosell, J., 2017. Quantitative and qualitative analysis for the study of Middle Paleolithic retouched artifacts: Unit III of Teixoneres cave (Barcelona, Spain). *Journal of Archaeological Science: Reports* 12, 658–672. <https://doi.org/10.1016/j.jasrep.2017.02.021>
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Machine learning* 20, 273–297.
- Courbin, P., Brenet, M., Michel, A., Gravina, B., 2020. Spatial analysis of the late Middle Palaeolithic open-air site of Bout-des-Vergnes (Bergerac, Dordogne) based on lithic technology and refitting. *Journal of Archaeological Science: Reports* 32, 102373. <https://doi.org/https://doi.org/10.1016/j.jasrep.2020.102373>
- Cover, T., Hart, P., 1967. Nearest neighbor pattern classification. *IEEE Transactions on Information Theory* 13, 21–27. <https://doi.org/10.1109/TIT.1967.1053964>
- Cramer, J.S., 2004. The early origins of the logit model. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 35, 613–626. <https://doi.org/10.1016/j.shpsc.2004.09.003>
- Delagnes, A., 1995. Variability within Uniformity: Three Levels of Variability within the Levallois System, in: Dibble, H.L., Bar-Yosef, O. (Eds.), *The Definition and Interpretation of Levallois Technology*, Monographs in World Archaeology. Prehistory Press, Madison, Wisconsin, pp. 201–211.
- Delagnes, A., Meignen, L., 2006. Diversity of Lithic Production Systems During the Middle Paleolithic in France. Are There Any Chronological Trends?, in: Hovers, E., Kuhn, S.L., Jochim, M. (Eds.), *Transitions Before the Transition Evolution and Stability in the Middle Paleolithic and Middle Stone Age*. Springer, pp. 85–107.
- Delpiano, D., Gennai, J., Peresani, M., 2021. Techno-Functional Implication on the Production of Discoid and Levallois Backed Implements. *Lithic Technology* 46, 171–191. <https://doi.org/10.1080/01977261.2021.1886487>
- Duran, J.-P., 2005. L'industrie moustérienne des ànecs (Rodès, Pyrénées-orientales, France). *PYRENAE* 36, 11–39.
- Duran, J.-P., Soler, N., 2006. Variabilité des modalités de débitage et des productions lithiques dans les

- industries moustériennes de la grotte de l'Arbreda, secteur alpha (Serinyà, Espagne). *Bulletin de la Société Préhistorique Française* 103, 241–262.
- Faivre, J.-Ph., Gravina, B., Bourguignon, L., Discamps, E., Turq, A., 2017. Late Middle Palaeolithic lithic technocomplexes (MIS 5–3) in the northeastern Aquitaine Basin: Advances and challenges. *Quaternary International* 433, 116–131. <https://doi.org/10.1016/j.quaint.2016.02.060>
- Fernandes, P., Morala, A., Schmidt, P., Séronie-Vivien, M.-R., Turq, A., 2012. Le silex du Bergeracois: État de la question. *Quaternaire Continental d'Aquitaine, excursion AFEQ, ASF 2012* 2012, 22–33.
- Fisher, R.A., 1936. The use of multiple measurements in taxonomic problems. *Annals of Eugenics* 7, 179–188.
- Folgado, M., Brenet, M., 2010. Economie de débitage et organisation de l'espace technique sur le site du Paléolithique moyen de plein-air de La Mouline (Dordogne, France), in: Conard, N., Delagnes, A. (Eds.), *Settlement Dynamics of the Middle Paleolithic and Middle Stone Age*. Kerns Verlag - (Tübingen Publications in Prehistory), Tübingen, pp. 427–454.
- Forestier, H., 1993. Le Clactonien: Mise en application d'une nouvelle méthode de débitage s'inscrivant dans la variabilité des systèmes de production lithique du Paléolithique ancien. *Paléo* 5, 53–82. <https://doi.org/10.3406/pal.1993.1104>
- Frey, P.W., Slate, D.J., 1991. Letter recognition using Holland-style adaptive classifiers. *Machine learning* 6, 161–182.
- Friedman, J.H., 2002. Stochastic gradient boosting. *Computational Statistics & Data Analysis* 38, 367–378. [https://doi.org/10.1016/S0167-9473\(01\)00065-2](https://doi.org/10.1016/S0167-9473(01)00065-2)
- Friedman, J.H., 2001. Greedy function approximation: A gradient boosting machine. *Annals of statistics* 29, 1189–1232.
- Geneste, J.-M., 1988. Les Industries De La Grotte Vaufrey: Technologie du débitage, économie et circulation de la matière première lithique, in: Rigaud, J.-P. (Ed.), *La Grotte Vaufrey à Cenac Et Saint-Julien (Dordogne) : Paléoenvirons, Chronologie Et Activités Humaines, Mémoires de La Société Préhistorique Française (Revue)*. Société préhistorique française, Paris, pp. 441–517.
- González-Molina, I., Jiménez-García, B., Mafllo-Fernández, J.-M., Baquedano, E., Domínguez-Rodrigo, M., 2020. Distinguishing Discoid and Centripetal Levallois methods through machine learning. *PLOS ONE* 15, e0244288. <https://doi.org/10.1371/journal.pone.0244288>
- Greenwell, B., Boehmke, B., Cunningham, J., Developers, G.B.M., Greenwell, M.B., 2019. Package ‘gbm’ R package version 2.
- Grimaldi, S., Santaniello, F., 2014. New insights into Final Mousterian lithic production in western Italy. *Quaternary International* 350, 116–129. <https://doi.org/10.1016/j.quaint.2014.03.057>
- Gunz, P., Mitteroecker, P., 2013. Semilandmarks: A method for quantifying curves and surfaces. *Hystrix* 24, 103–109. <https://doi.org/10.4404/hystrix-24.1-6292>
- Gunz, P., Mitteroecker, P., Bookstein, F.L., 2005. Semilandmarks in three dimensions, in: *Modern Morphometrics in Physical Anthropology*. Springer, New York, pp. 73–98.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An Introduction to Statistical Learning with Applications in R*, Second Edition. ed. Springer.
- Karatzoglou, A., Meyer, D., Hornik, K., 2006. Support Vector Machines in R. *Journal of Statistical Software* 15, 1–28. <https://doi.org/10.18637/jss.v015.i09>
- Karatzoglou, A., Smola, A., Hornik, K., Zeileis, A., 2004. Kernlab - An S4 Package for Kernel Methods in R. *Journal of Statistical Software* 11, 1–20. <https://doi.org/10.18637/jss.v011.i09>
- Kelly, H., 1954. Contribution à l'étude de la technique de la taille levalloisienne. *Bulletin de la Société Préhistorique Française* 51, 149–169. <https://doi.org/10.3406/bspf.1954.3077>
- Kendall, D.G., 1984. Shape Manifolds, Procrustean Metrics, and Complex Projective Spaces. *Bulletin of the London Mathematical Society* 16, 81–121. <https://doi.org/10.1112/blms/16.2.81>
- Kuhn, M., 2008. Building Predictive Models in R using the caret Package. *Journal of Statistical Software* 28. <https://doi.org/10.18637/jss.v028.i05>
- Kuhn, S.L., 2013. Roots of the Middle Paleolithic in Eurasia. *Current Anthropology* 54, S255–S268. <https://doi.org/10.1086/673529>
- Lantz, B., 2019. Machine learning with R: Expert techniques for predictive modeling. Packt publishing ltd.
- Lenoir, M., Turq, A., 1995. Recurrent Centripetal Debitage (Levallois and Discoidal): Continuity or Discontinuity?, in: Dibble, H.L., Bar-Yosef, O. (Eds.), *The Definition and Interpretation of Levallois Technology*,

- Monographs in World Archaeology. Prehistory Press, Madison, Wisconsin, pp. 249–256.
- Locht, J.-L., 2003. L'industrie lithique du gisement de Beauvais (Oise, France): Objectifs et variabilité du débitage discoïde, in: Peresani, M. (Ed.), Discoid Lithic Technology: Advances and Implications, BAR International Series. Archaeopress, Oxford, pp. 193–209.
- Marciani, G., Ronchitelli, A., Arrighi, S., Badino, F., Bortolini, E., Boscato, P., Boschin, F., Crezzini, J., Delpiano, D., Falcucci, A., Figus, C., Lugli, F., Oxilia, G., Romandini, M., Riel-Salvatore, J., Negrino, F., Peresani, M., Spinapolic, E.E., Moroni, A., Benazzi, S., 2020. Lithic techno-complexes in Italy from 50 to 39 thousand years BP: An overview of lithic technological changes across the Middle-Upper Palaeolithic boundary. *Quaternary International* 551, 123–149. <https://doi.org/10.1016/j.quaint.2019.11.005>
- Martín-Viveros, J.I., Ollé, A., Chacón, M.G., Romagnoli, F., Gómez de Soler, B., Vaquero, M., Saladié, P., Vallverdú, J., Carbonell, E., 2020. Use-wear analysis of a specific mobile toolkit from the Middle Paleolithic site of Abric Romaní (Barcelona, Spain): A case study from level M. *Archaeological and Anthropological Sciences* 12, 16. <https://doi.org/10.1007/s12520-019-00951-z>
- McPherron, S., 2019. E5 (Beta Version).
- Meignen, L., 1996. Persistance des traditions techniques dans l'abri des Canalettes (Nant-Aveyron). *Quaternaria Nova* 6, 449–64.
- Meignen, L., 1993. Les industries lithiques de l'abri des Canalettes: Cuche 2, in: Meignen, L. (Ed.), L'abri Des Canalettes. Un Habitat Moustérien Sur Les Grands Causses (Nant-Aveyron), Monographie Du CRA. CNRS Ed., Paris, pp. 238–328.
- Mitteroecker, P., Gunz, P., 2009. Advances in Geometric Morphometrics. *Evolutionary Biology* 36, 235–247. <https://doi.org/10.1007/s11692-009-9055-x>
- Mourre, V., 2003. Discoïde ou pas Discoïde? Réflexions sur la pertinence des critères techniques définissant le débitage discoïde, in: Peresani, M. (Ed.), Discoid Lithic Technology. Advances and Implications, BAR International Series. Archaeopress, Oxford, pp. 1–17.
- O'Higgins, P., 2000. The study of morphological variation in the hominid fossil record: Biology, landmarks and geometry. *Journal of Anatomy* 197, 103–120. <https://doi.org/10.1046/j.1469-7580.2000.19710103.x>
- Pasty, J.-F., Liegard, S., Alix, P., 2004. Étude de l'industrie lithique du site paléolithique moyen des Fendeux (Coulanges, Allier). *Bulletin de la Société préhistorique française* 101, 5–25. <https://doi.org/10.3406/bspf.2004.12945>
- Pearson, K., 1901. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 2, 559–572. <https://doi.org/10.1080/14786440109462720>
- Picin, A., Chacón, M.G., Gómez de Soler, B., Blasco, R., Rivals, F., Rosell, J., 2020. Neanderthal mobile toolkit in short-term occupations at Teixoneres Cave (Moia, Spain). *Journal of Archaeological Science: Reports* 29, 102165. <https://doi.org/10.1016/j.jasrep.2019.102165>
- Quinlan, J.R., 2014. C4. 5: Programs for machine learning. Elsevier.
- Quinlan, J.R., 1996. Improved Use of Continuous Attributes in C4.5. *Journal of Artificial Intelligence Research* 4, 77–90. <https://doi.org/10.1613/jair.279>
- Ridgeway, G., 2007. Generalized Boosted Models: A guide to the gbm package. R package vignette 2007.
- Rios-Garaizar, J., 2010. Organización económica de las sociedades neandertales: El caso del Nivel VIII de Amalda (Guipuzkoa). *Zephyrus* LXV, 15–37.
- Ríos-Garaizar, J., 2017. A new chronological and technological synthesis for Late Middle Paleolithic of the Eastern Cantabrian Region. *Quaternary International* 433, 50–63. <https://doi.org/10.1016/j.quaint.2016.02.020>
- Roebroeks, W., Loecker, D.D., Hennekens, P., Leperen, M.V., 1992. "A veil of stones": On the interpretation of an early Middle Palaeolithic low density scatter at Maastricht-Belvédère (The Netherlands). *Analecta Praehistorica Leidensia* 25| The end of our third decade: Papers written on the occasion of the 30th anniversary of the Institutte of prehistory, volume I 25, 1–16.
- Romagnoli, Francesca, Bargalló, A., Chacón, M.G., Gómez de Soler, B., Vaquero, M., 2016. Testing a hypothesis about the importance of the quality of raw material on technological changes at Abric Romaní (Capellades, Spain): Some considerations using a high-resolution techno-economic perspective. *Journal of Lithic Studies* 3, 635–659. <https://doi.org/10.2218/jls.v3i2.1443>
- Romagnoli, F., Trenti, F., Nannini, L., Carmignani, L., Ricci, G., Lo Vetro, D., Martini, F., Sarti, L., 2016. Raw material procurement and productive sequences in the Palaeolithic of southern Italy: The Tyrrenian

- and the Ionian areas. Ressources lithiques, productions et transferts entre Alpes et Méditerranée 185–205.
- Rumelhart, D.E., Hinton, G.E., Williams, R.J., 1986. Learning representations by back-propagating errors. *Nature* 323, 533–536.
- Schlager, S., 2017. Morpho and Rvcg–Shape Analysis in R: R-Packages for geometric morphometrics, shape analysis and surface manipulations, in: Statistical Shape and Deformation Analysis. Elsevier, pp. 217–256.
- Shott, M.J., 2018. The Costs and Benefits of Technological Organization: Hunter-Gatherer Lithic Industries and Beyond, in: Robinson, E., Sellet, F. (Eds.), *Lithic Technological Organization and Paleoenvironmental Change*. Springer International Publishing, Cham, pp. 321–333. [https://doi.org/10.1007/978-3-319-64407-3\\_15](https://doi.org/10.1007/978-3-319-64407-3_15)
- Simpson, E.H., 1951. The Interpretation of Interaction in Contingency Tables. *Journal of the Royal Statistical Society: Series B (Methodological)* 13, 238–241. <https://doi.org/10.1111/j.2517-6161.1951.tb00088.x>
- Slimak, L., 2003. Les Débitages discoïdes moustériens: Evaluation d'un concept technologique, in: Peresani, M. (Ed.), *Discoid Lithic Technology. Advances and Implications*, BAR International Series. Archaeopress, Oxford, pp. 33–65.
- Slimak, L., 1998. La variabilité des débitages discoïdes au Paléolithique moyen: Diversité des méthodes et unité d'un concept. L'exemple des gisements de la Baume Néron (Soyons, Ardèche) et du Champ Grand (Saint-Maurice-sur-Loire, Loire). *Préhistoire anthropologie méditerranéennes* 7, 75–88.
- Soressi, M., Geneste, J.-M., 2011. The History and Efficacy of the Chaîne Opératoire Approach to Lithic Analysis: Studying Techniques to Reveal Past Societies in an Evolutionary Perspective. *PaleoAnthropology* 2011, 334–350. <https://doi.org/10.4207/PA.2011.ART63>
- Spackman, K.A., 1989. Signal detection theory: Valuable tools for evaluating inductive learning, in: Proceedings of the Sixth International Workshop on Machine Learning. Elsevier, pp. 160–163.
- Team, R.C., 2019. R: A language and environment for statistical computing.
- Team, R.s., 2019. RStudio: Integrated Development for R.
- Turq, A., Roebroeks, W., Bourguignon, L., Faivre, G.-P., 2013. The fragmented character of Middle Palaeolithic stone tool technology. *Journal of Human Evolution* 65, 641–655. <https://doi.org/10.1016/j.jhevol.2013.07.014>
- Walker, S.H., Duncan, D.B., 1967. Estimation of the Probability of an Event as a Function of Several Independent Variables. *Biometrika* 54, 167–179. <https://doi.org/10.2307/2333860>
- Weih, C., Ligges, U., Luebke, K., Raabe, N., 2005. klaR analyzing German business cycles, in: Data Analysis and Decision Support. Springer, pp. 335–343.
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., Kuhn, M., Pedersen, T., Miller, E., Bache, S., Müller, K., Ooms, J., Robinson, D., Seidel, D., Spinu, V., Takahashi, K., Vaughan, D., Wilke, C., Woo, K., Yutani, H., 2019. Welcome to the Tidyverse. *Journal of Open Source Software* 4, 1686. <https://doi.org/10.21105/joss.01686>
- Wright, M.N., Ziegler, A., 2017. Ranger: A Fast Implementation of Random Forests for High Dimensional Data in C++ and R. *Journal of Statistical Software* 77, 1–17. <https://doi.org/10.18637/jss.v077.i01>