

# Paper Fingerprint by Forming Fabric: A Univariate Feature Selection Approach Using Periodic Marks Analysis

Yong Ju Lee,<sup>a</sup> Chang Woo Jeong,<sup>b,\*</sup> Tai-Ju Lee,<sup>a</sup> Geon-Woo Kim,<sup>a</sup> and Hyoung Jin Kim<sup>a,\*</sup>

Evidence by which to confirm the location and approximate manufacturing date of document paper is a critical task in forensic investigations, particularly in cases involving suspected forgery or document manipulation. In this study, periodic marks formed during the papermaking process were analyzed using light-transmitted images captured by a two-dimensional lab formation sensor. Step and angle data from the top five intensity peaks were extracted and used to train tree-based classification models. To handle directional symmetry, a modulo 180° transformation was applied to the angle data. The random forest (RF) classifier outperformed decision tree (DT) and extreme gradient boosting (XGB) models, achieving the highest F1 score. Feature importance analysis revealed that the step and angle at the third intensity level were the most discriminative features, likely reflecting structural characteristics of forming fabrics or drainage patterns. A simplified univariate strategy using these features also showed potential for estimating production periods. However, differences between the top and bottom surfaces—particularly in twin-wire systems—introduced classification bias, highlighting the need to separately classify paper sides in forensic datasets. Overall, this study demonstrates the feasibility and limitations of using periodic mark analysis for document dating.

DOI: 10.15376/biores.20.4.9208-9225

**Keywords:** Random forest (RF); Feature importance; DBSCAN; Forming fabric; Light-transmitted image

**Contact information:** a: Department of Forest Products and Biotechnology, Kookmin University, 77 Jeongneung-ro, Seongbuk-gu, Seoul 02707 Republic of Korea; b: Graduate School of Scientific Criminal Investigation, Chungnam National University, Daejeon 34134, Korea;

\* Corresponding authors: [hyjikim@kookmin.ac.kr](mailto:hyjikim@kookmin.ac.kr); [leetj@kookmin.ac.kr](mailto:leetj@kookmin.ac.kr)

## INTRODUCTION

Evidence related to manufacturing dates and locations is essential for detecting added or substituted pages in forged documents. Determining the date of a document helps establish the period during which a tax document or other official document was produced. This is performed by comparing the date of the document with the earliest availability of the paper on which it is printed. The materials used in the document must have been available before or on the date of the document; otherwise, if the date of the document is earlier than the availability of its paper, this anachronism indicates backdating (Ellen *et al.* 2018). To facilitate this analysis, an accessible reference database is required (Jeong *et al.* 2024a, b).

Conventional paper analysis methods focus on evaluating physical and optical characteristics, such as tensile strength, thickness, basis weight, ash content, color, and fluorescence (Grant 1973). In addition, techniques such as X-ray diffraction (Foner and Adan 1983), elemental analysis (Spence *et al.* 2002), infrared spectroscopy (Hwang *et al.* 2024; Lee *et al.* 2024a, b; Lee *et al.* 2025a), Raman spectroscopy (Kuptsov 1994), image analysis (Miyata *et al.* 2002), and pyrolysis gas chromatography (Ebara *et al.* 1982) have been applied. However, these methods typically face challenges in accurately identifying the manufacturing date of paper. For example, physical testing methods are limited by their destructive nature, which may damage the documents. Nondestructive techniques, such as spectroscopy analysis, also have limitations, for example, documents may be contaminated, and the aging characteristics of paper can change depending on storage conditions (Havlíková *et al.* 2009; Małachowska *et al.* 2021).

To address these limitations, light-transmitted image analysis methods based on fast Fourier transformation have been developed (Miyata *et al.* 2002; Comte *et al.* 2006; Berger and Ramos 2012). Berger (2009) captured light-transmitted images of office paper and then performed frequency analysis using the two-dimensional (2D) power spectrum technique. This method measures similarity through correlation and shows promise in distinguishing common office papers. Jeong *et al.* (2024a,b) performed classification analysis to discriminate manufacturers using periodic marks from the forming fabrics used during the papermaking process. Light-transmitted images were captured using a 2D lab formation sensor (Techpap, France), and the images were then transformed into Fourier transform-based images. From these images, periodic marks were derived from the weaving patterns of the forming fabric and drainage marks depending on the paper machines.

These periodic patterns primarily originate from the forming fabric, an endless woven belt composed of synthetic monofilaments. It has two distinct sides—a forming surface and a wearing surface—with typical filament spacings of 0.3 to 0.6 mm (Adanur 2017). During the paper formation process, the fiber suspension (headbox jet) is delivered onto the forming surface, where water drainage and fiber deposition inevitably imprint the mesh structure onto the paper surface. These imprints are known as weave marks or wire marks. In addition to weave marks, other periodic marks arise throughout the papermaking process due to mechanical interactions in various machine sections. These include: (1) drainage marks from variations in dewatering in the wire section, (2) shadow marks from suction rolls, (3) press felt marks from the press section, (4) groove marks from shoe press belts, and (5) canvas and dryer fabric marks from the drying section. Because such periodic marks are consistently and uniquely introduced during paper manufacturing, they serve as valuable forensic features. They have been widely documented for applications in paper origin discrimination and manufacturing date estimation (Berger 2009; Jeong *et al.* 2024a, b; Lee *et al.* 2024c; Lee *et al.* 2025b).

Distinguishing paper manufacturers and products can support forensic investigations involving document alterations, such as detecting substituted pages or inserted content in multipage documents (Berger and Ramos 2012; Comte *et al.* 2006; Miyata *et al.* 2002). While various analytical techniques for paper source identification have been extensively reported, these approaches do not offer the same level of forensic resolution as methods aimed at determining the manufacturing date of a document. In contrast, studies focused on document dating remain relatively limited, primarily due to the inherent challenges associated with such analyses. These challenges include potential contamination, as well as the fact that the aging behavior of paper is highly sensitive to storage conditions, making it difficult to establish consistent chronological markers.

In the authors' previous studies on forensic document dating, periodic marks were analyzed using a 2D lab formation sensor (Lee *et al.* 2024c). To the best of the authors' knowledge, periodic marks and ash content are intrinsic characteristics of paper that remain stable under environmental exposure and are resistant to post-manufacturing tampering, thereby serving as robust forensic markers (Choi *et al.* 2018). Nevertheless, ash content analysis often requires a relatively large sample size and may involve destructive procedures (Choi *et al.* 2018). In the authors' earlier work, an artificial neural network (ANN) classifier was developed using features derived from periodic surface patterns, achieving a high classification performance with an F1 score of 0.951.

Although this analytical approach is highly effective, ANN models are inherently prone to overfitting and are known to exhibit limited extrapolation capabilities (Rhein *et al.* 2024). Furthermore, neural networks often function as "black box" models, offering little interpretability with respect to the meaning or significance of individual weights (Lundberg and Lee 2017). Most importantly, machine learning-based training and prediction processes are often not user-friendly, requiring specialized expertise and computational resources that may limit their accessibility in practical forensic settings.

The objective of this study was to develop user-friendly methods, such as univariate approaches, for paper differentiation. Previous studies have primarily employed multivariate and ANN-based classification models. In this study, a 2D lab formation sensor was used to extract features related to periodic marks on the paper surface. Feature importance was then evaluated using a random forest (RF) model to identify the most relevant periodic mark characteristics. Ultimately, the aim is to identify key periodic mark features that are critical for forensic document dating and to propose a method for constructing a database that integrates these features using an accessible, user-friendly analytical approach.

## EXPERIMENTAL

### Materials

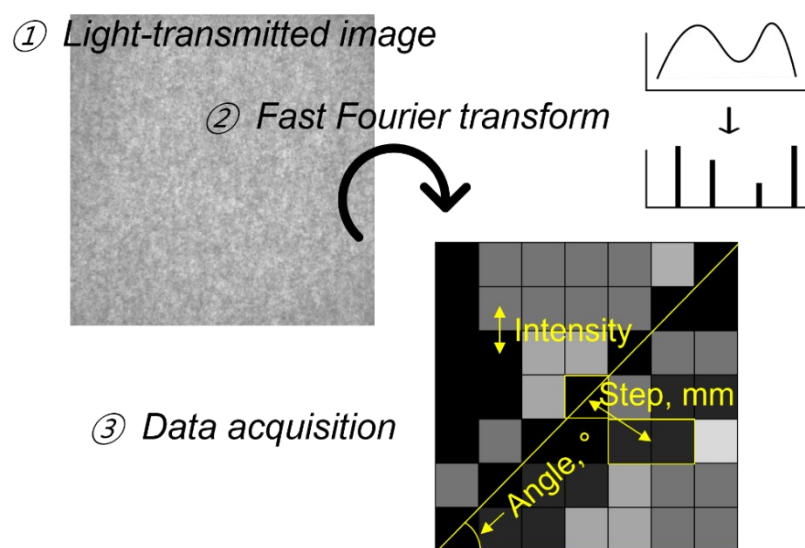
The information on the printing paper used in this study is presented in Table 1. For document paper dating, 12 products were collected, all from the same manufacturer but differing by production date. The "H" product was typical office paper in A4 size (210 mm × 297 mm).

**Table 1.** Information on Printing Paper with Different Manufacturing Dates

| NO. | Product                                      | Code    | Manufacturing Date          | Grammage (g/m <sup>2</sup> ) |
|-----|----------------------------------------------|---------|-----------------------------|------------------------------|
| 1   | "H"<br>product<br>from a<br>Paper<br>Company | H202009 | Sep 19 <sup>th</sup> , 2020 | 80                           |
| 2   |                                              | H202012 | Dec 13 <sup>th</sup> , 2020 |                              |
| 3   |                                              | H202103 | Mar 01 <sup>st</sup> , 2021 |                              |
| 4   |                                              | H202106 | Jun 02 <sup>nd</sup> , 2021 |                              |
| 5   |                                              | H202109 | Sep 07 <sup>th</sup> , 2021 |                              |
| 6   |                                              | H202112 | Dec 17 <sup>th</sup> , 2021 |                              |
| 7   |                                              | H202204 | Apr 27 <sup>th</sup> , 2022 |                              |
| 8   |                                              | H202206 | Jun 30 <sup>th</sup> , 2022 |                              |
| 9   |                                              | H202210 | Oct 18 <sup>th</sup> , 2022 |                              |
| 10  |                                              | H202301 | Jan 28 <sup>th</sup> , 2023 |                              |
| 11  |                                              | H202304 | Apr 10 <sup>th</sup> , 2023 |                              |
| 12  |                                              | H202307 | Jul 19 <sup>th</sup> , 2023 |                              |

## Dataset

A 2D lab formation sensor (2D-F sensor) was used to scan the printing paper. The data collection process was repeated, with 50 samples collected for each product. In addition, measurements were performed randomly without distinguishing between the top and bottom sides. The 2D-F sensor generated light-transmitted images automatically, which were then analyzed using a fast Fourier transform algorithm to examine three distinct factors in the images, *i.e.*, intensity, angle, and step (Jeong *et al.* 2024a,b; Lee *et al.* 2024c; Lee *et al.* 2025b). Figure 1 illustrates the principles for calculating the intensity, step, and angle from the light-transmitted images. The intensity refers to the brightness of periodic marks on paper. The term “angle” indicates the orientation of the linear repetitive marks within the plane of the paper, while “step” refers to the spacing between these marks, measured in millimeters.



**Fig. 1.** Principles for calculating intensity, step, and angle from light-transmitted images

The step and angle data of the periodic marks were measured at the highest 10 intensities, creating a  $1 \times 30$  array from each measurement. However, the authors noted that the intensity data varied for each measurement. For this reason, intensity values were removed from the dataset. The data scanned from a single paper product were represented as a 50 (samples)  $\times$  20 (variables) matrix, resulting in 600 (50 (samples)  $\times$  12 (products)) data points.

## Data Preprocessing

Angles can be considered symmetric within the same step (*e.g.*,  $60^\circ$  and  $240^\circ$ ), and they are treated as equivalent. Therefore, all angle data were converted by modulo 180 transformation, as defined in Eq. 1.

$$\theta' = \theta \bmod 180 \quad (1)$$

The usefulness of such angular transformations for capturing the directional characteristics of periodic marks has already been demonstrated in prior studies (Jeong *et al.* 2024a,b; Lee *et al.* 2024c), where the absolute value of the sine function,  $\sin \theta$ , was used to account for symmetry.

## PCA of Step and Angle Data

Principal component analysis (PCA) was performed to reveal the inherent patterns and hidden features in the high-dimensional dataset. The step and angle data of the periodic marks on the paper comprised 20-dimensional data. The 20-dimensional data were projected onto a new seven-dimensional orthogonal space. Finally, the data structure was visualized in 2D space. The usefulness of PCA based on features of periodic marks on paper was demonstrated in our previous research (Lee *et al.* 2025b). However, that study relied on intensity data, which can be variable, to differentiate the paper origin.

Before performing PCA, the step and angle data were standardized using Z-score normalization to scale each variable to have a mean of 0 and a variance of 1. The scaled value  $Z_i$  is defined as in Eq. 2,

$$Z_i = \frac{x_i - \mu}{\sigma} \quad (2)$$

where  $x_i$  denotes each observation of a variable  $x$ ,  $\mu$  denotes the mean of the variable  $x$ , and  $\sigma$  denotes the standard deviation of the variable  $x$ .

## Outlier Detection in Dataset using DBSCAN

The density-based spatial clustering of applications with noise (DBSCAN) algorithm was used to detect outliers in the dataset (Ester *et al.* 1996; Campello *et al.* 2013), which were projected onto the orthogonal coordinate system of principal components (PCs). The DBSCAN parameters epsilon (eps) and minimum points (minPts) were empirically set to 5.0 and 7, respectively. Here, “eps” represents the radius within which neighboring data points exert mutual influence, and “minPts” specifies the minimum number of data points required to form a cluster. A cluster was established when at least seven consecutive points were found within a distance of 5.0 from a given data point.

The detected outliers were removed from the dataset, and the classification model was then trained on the imbalanced dataset.

## Classification Models for Forensic Document Dating

For the classification modeling of forensic document dating, tree-based models were employed. Tree-based models are widely recognized for providing a robust framework for decision-making processes. In this study, two representative ensemble methods—random forest (RF) and extreme gradient boosting (XGBoost)—were constructed, both of which utilize decision trees (DTs) as base learners.

All data processing and classification were carried out in R (R Core Team 2023, version 4.5.0, Auckland, New Zealand).

### *Dataset splitting*

The dataset was divided into training and test sets at a ratio of 7:3 for training and evaluation of the classification models to be compared. Stratified random sampling was used to ensure that the split ratio was maintained across all classes.

### *Decision tree (DT)*

Single DTs (Breiman 2017) were tested to compare their classification performance with that of the RF model, which is an ensemble learning method. To optimize the DT model, the complexity parameter (cp), which controls the tree size and complexity, was carefully tuned. A grid search within the range of 0.001 to 0.1 was performed to identify the optimal cp value, thereby balancing model accuracy and preventing overfitting. The

final model was selected based on the cp value that achieved the highest performance in a threefold cross-validation.

#### *Random forest (RF)*

The RF classifier (Breiman 2001), an ensemble learning technique, helps mitigate premature convergence to some extent. Ensemble learning combines predictions from multiple models to improve the accuracy beyond that achievable by individual models. In this study, DTs (Breiman 2017) were employed as the base learners to construct the RF model. To increase the diversity among the DTs, the RF model builds them through random subsampling, without using all input variables, to construct independent trees. Furthermore, the RF model generates bootstrap samples from the training dataset via random sampling with replacement. In this process, approximately two-thirds of the training data (in-bag samples) are used to train the DTs, and the remaining one-third (out-of-bag (OOB) samples) are used to validate the RF model. The probability that a sample is not selected from a dataset of size  $m$  during random sampling with replacement is represented by  $(m-1)/m$ . When repeated  $m$  times, the OOB probability approaches 36.8%, as shown in Eq. 3.

$$OOB = \lim_{m \rightarrow \infty} \left(1 - \frac{1}{m}\right)^m = e^{-1} \approx 0.3678. \quad (3)$$

For the input variable `n_feature` used in DT generation, the square root (“sqrt”), binary logarithm (“log2”), and one-third (“1/3”) of all spectral points were tested. The number of trees (`n_tree`) was varied from 10 to 150, and the optimal values for `n_feature` and `n_tree` were selected based on the minimum OOB error identified through the grid search.

#### *XGBoost (XGB)*

The XGB algorithm (Chen and Guestrin 2016) is a tree-based ensemble learning method that differs fundamentally from the RF approach. While RF builds multiple DTs independently and aggregates their outputs through majority voting, XGB constructs DTs sequentially, where each tree attempts to correct the residual errors of the previous trees through gradient-based optimization. This boosting framework allows XGB to achieve higher prediction accuracy, especially in cases involving complex, nonlinear patterns.

In this study, DTs were employed as base learners to construct the XGB model. To enhance generalization performance and prevent overfitting, XGB incorporates several regularization techniques, including shrinkage (learning rate control), subsampling of rows and columns, and penalties on tree complexity. Additionally, unlike traditional gradient boosting, XGB leverages second-order derivatives (*i.e.*, both gradients and Hessians) of the loss function to improve convergence speed and model stability.

For multiclass classification, the number of boosting rounds (`n_rounds`) was varied from 10 to 150, the maximum tree depth (`max_depth`) was adjusted between 3 and 5, and the learning rate (`eta`) was tested in the range of 0.001 to 0.3. Grid search was used to identify the optimal set of hyperparameters based on the classification accuracy on the validation dataset.



### Evaluation metric

Accurately evaluating the classification of observations into positive and negative categories is essential in classification tasks. True positives represent observations correctly classified as belonging to the positive class, whereas true negatives denote observations correctly classified as belonging to the negative class. Conversely, false negatives occur when positive observations are incorrectly classified as negative, and false positives arise when negative observations are incorrectly classified as positive. These values enable the calculation of various performance metrics to assess the effectiveness of classification models in detecting the target class.

The F1 score, which is the harmonic mean of precision (Eq. 4) and recall (Eq. 5), is a commonly used performance metric in classification problems with class imbalance; it is defined in Eq. 6.

$$Precision = \frac{TP}{TP+FP}, \quad (4)$$

$$Recall = \frac{TP}{TP+FN}, \quad (5)$$

$$F1\ score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

The weighted F1 score, which was applied to evaluate the performance of the RF model, accounts for class imbalance by assigning weights to each class (Eq. 7) and integrating these weights into their respective F1 scores (Eq. 8). This approach enables the weighted F1 score to assess both individual class performance and overall model performance, even with imbalanced datasets (Hwang *et al.* 2024),

$$w_i = \frac{N_i}{T_i}, \quad (7)$$

where  $w_i$  denotes the weight of class,  $N_i$  denotes the number of samples in the class, and  $T_i$  denotes the total number of samples.

$$Weighted\ F1_i = \sum_{i=1}^N w_i \times F1_i. \quad (8)$$

### Feature Importance Measures from the Random Forest Model

The importance of order in the step and angle data was assessed using the mean decrease impurity method (Louppe *et al.* 2013). This approach identifies variables that are significant in the decision making of the RF model to confirm the manufacturing date and location of printing paper. The DTs contribute to understanding the role of each variable and offer transparency and insight into the decision-making process of the classification model. This process, called feature importance measurement, is defined as follows,

$$I(n_j) = w_j C_j - w_{L_j} C_{L_j} - w_{R_j} C_{R_j}, \quad (9)$$

where  $n_j$  denotes the parent node,  $L_j$  and  $R_j$  denote the left and right child nodes branched from  $n_j$ ,  $w_j$  denotes the node weight or sample count, and  $C_j$  denotes the impurity of  $n_j$ . The importance of variable  $i$  in a DT is defined as follows,

$$I(f_i) = \frac{\sum_j I(n_j)}{\sum_{k \in n_{all}} I(n_k)} \quad (10)$$

where  $I(f_i)$  denotes the importance of variable  $i$  in the tree model. In the RF model, variable importance is determined by aggregating the importance values of all DTs. Prior

to aggregation, the importance of each variable is normalized using Eq. 11.

$$\text{normI}(f_i) = \frac{I(f_i)}{\sum_{j \in f_{all}} I(f_j)} \quad (11)$$

Finally, the overall importance of each variable in the RF model is averaged across all DTs, as shown in Eq. 12,

$$I(RF_i) = \frac{\sum_{j \in t_{all}} \text{normI}(f_{ij})}{T} \quad (12)$$

where  $t$  represents each DT model,  $\text{normI}(f_{ij})$  denotes the normalized importance of variable  $i$  in the RF model, and  $T$  represents the total number of DTs. The selected features related to periodic marks, such as their angle and step (Fig. 1), were employed to develop a simple method for printing paper dating.

## RESULTS AND DISCUSSION

### Step and Angle of Periodic Marks on Printing Paper

Tables 2 and 3 present the step and angle data of the periodic marks on paper from a single measurement. For the step and angle data at the first two intensity levels, no differences were observed across all classes, because these marks were derived from the weaving patterns. The weave marks in the machine direction and the cross direction of the paper machine are common features found in all types of products, regardless of the manufacturer or manufacturing date. Angles of 180°, 360°, and 90° were considered in these cases. However, unique steps and angles, particularly at the third intensity level, depend on the forming fabrics used or the papermaking process (Jeong *et al.* 2024a,b; Lee *et al.* 2024c). For example, a step of 1.60 mm and an angle of 132.5° represent the distinctive characteristics of sample H202103. Therefore, these features are considered significant for forensic document examination. From this perspective, it can be hypothesized that if document paper can be differentiated based on the step and angle values at three, four, or five specific intensity levels, then a univariate method could be developed. In other words, paper samples may be distinguishable by simply comparing the step and angle at selected intensity levels.

**Table 2.** Steps (mm) of Periodic Marks on Paper at the Top Ten Intensities

| No. | Step 1 | Step 2 | Step 3 | Step 4 | Step 5 | Step 6 | Step 7 | Step 8 | Step 9 | Step 10 |
|-----|--------|--------|--------|--------|--------|--------|--------|--------|--------|---------|
| 1   | 1.33   | 1.33   | 1.73   | 1.64   | 3.27   | 2.90   | 2.72   | 3.18   | 2.00   | 2.35    |
| 2   | 1.32   | 1.32   | 1.70   | 1.72   | 1.12   | 1.64   | 3.34   | 3.29   | 2.85   | 2.60    |
| 3   | 1.32   | 1.32   | 1.60   | 1.75   | 1.26   | 2.86   | 2.24   | 2.94   | 1.84   | 2.14    |
| 4   | 1.31   | 1.31   | 1.61   | 3.21   | 3.04   | 1.73   | 3.13   | 2.90   | 2.25   | 3.06    |
| 5   | 1.32   | 1.32   | 1.09   | 1.27   | 3.29   | 2.18   | 2.80   | 1.53   | 1.38   | 3.01    |
| 6   | 1.32   | 1.32   | 1.66   | 1.25   | 1.13   | 3.30   | 3.12   | 3.18   | 2.59   | 2.68    |
| 7   | 1.32   | 1.32   | 1.56   | 2.90   | 3.26   | 3.32   | 2.18   | 2.76   | 2.82   | 3.07    |
| 8   | 1.31   | 1.31   | 1.66   | 1.73   | 1.11   | 3.09   | 2.29   | 1.57   | 2.24   | 1.77    |
| 9   | 1.31   | 1.32   | 1.76   | 3.07   | 1.21   | 2.41   | 2.58   | 2.47   | 1.13   | 3.02    |
| 10  | 1.32   | 1.32   | 3.26   | 2.90   | 2.83   | 2.14   | 2.48   | 1.61   | 2.31   | 2.50    |
| 11  | 1.31   | 1.31   | 1.64   | 3.16   | 3.17   | 1.89   | 2.35   | 2.94   | 2.94   | 2.61    |
| 12  | 1.32   | 1.32   | 1.64   | 1.68   | 3.29   | 2.94   | 3.21   | 3.10   | 2.88   | 1.60    |

1, H202009; 2, H202012; 3, H202103; 4, H202106; 5, H202109; 6, 202112; 7, H202204; 8, H202206; 9, H202210; 10, H202301; 11, H202304; 12, H202307



**Table 3.** Angles (°) of Periodic Marks on Paper at the Top Ten Intensities

| No. | Angle 1 | Angle 2 | Angle 3 | Angle 4 | Angle 5 | Angle 6 | Angle 7 | Angle 8 | Angle 9 | Angle 10 |
|-----|---------|---------|---------|---------|---------|---------|---------|---------|---------|----------|
| 1   | 179.4   | 360.0   | 52.7    | 130.8   | 94.4    | 174.8   | 128.1   | 23.4    | 170.9   | 166.2    |
| 2   | 179.4   | 360.0   | 130.7   | 49.9    | 90.0    | 175.6   | 172.5   | 29.2    | 129.8   | 120.6    |
| 3   | 179.4   | 360.0   | 132.5   | 53.9    | 100.8   | 79.7    | 19.4    | 147.4   | 115.3   | 103.5    |
| 4   | 179.4   | 360.0   | 131.9   | 162.5   | 72.8    | 52.7    | 94.0    | 142.4   | 74.7    | 92.2     |
| 5   | 179.4   | 360.0   | 95.4    | 100.9   | 18.0    | 115.3   | 118.8   | 117.5   | 103.9   | 66.5     |
| 6   | 179.4   | 360.0   | 132.4   | 87.8    | 93.0    | 168.1   | 54.3    | 35.2    | 31.8    | 139.7    |
| 7   | 179.4   | 360.0   | 133.0   | 5.2     | 37.8    | 58.8    | 148.2   | 36.8    | 68.3    | 110.8    |
| 8   | 179.4   | 360.0   | 47.6    | 128.4   | 87.0    | 15.4    | 104.5   | 58.1    | 94.2    | 93.6     |
| 9   | 179.4   | 360.0   | 55.7    | 120.3   | 102.6   | 142.3   | 75.8    | 109.9   | 94.3    | 6.1      |
| 10  | 179.4   | 360.0   | 127.8   | 95.2    | 55.2    | 122.8   | 89.0    | 63.1    | 110.7   | 174.4    |
| 11  | 179.4   | 360.0   | 129.8   | 40.0    | 132.0   | 90.9    | 105.3   | 105.4   | 56.3    | 115.7    |
| 12  | 179.4   | 360.0   | 129.8   | 51.9    | 72.0    | 99.3    | 78.1    | 102.8   | 148.8   | 96.6     |

1, H202009; 2, H202012; 3, H202103; 4, H202106; 5, H202109; 6, 202112; 7, H202204; 8, H202206; 9, H202210; 10, H202301; 11, H202304; 12, H202307

### PCA and Outlier Detection of Step and Angle

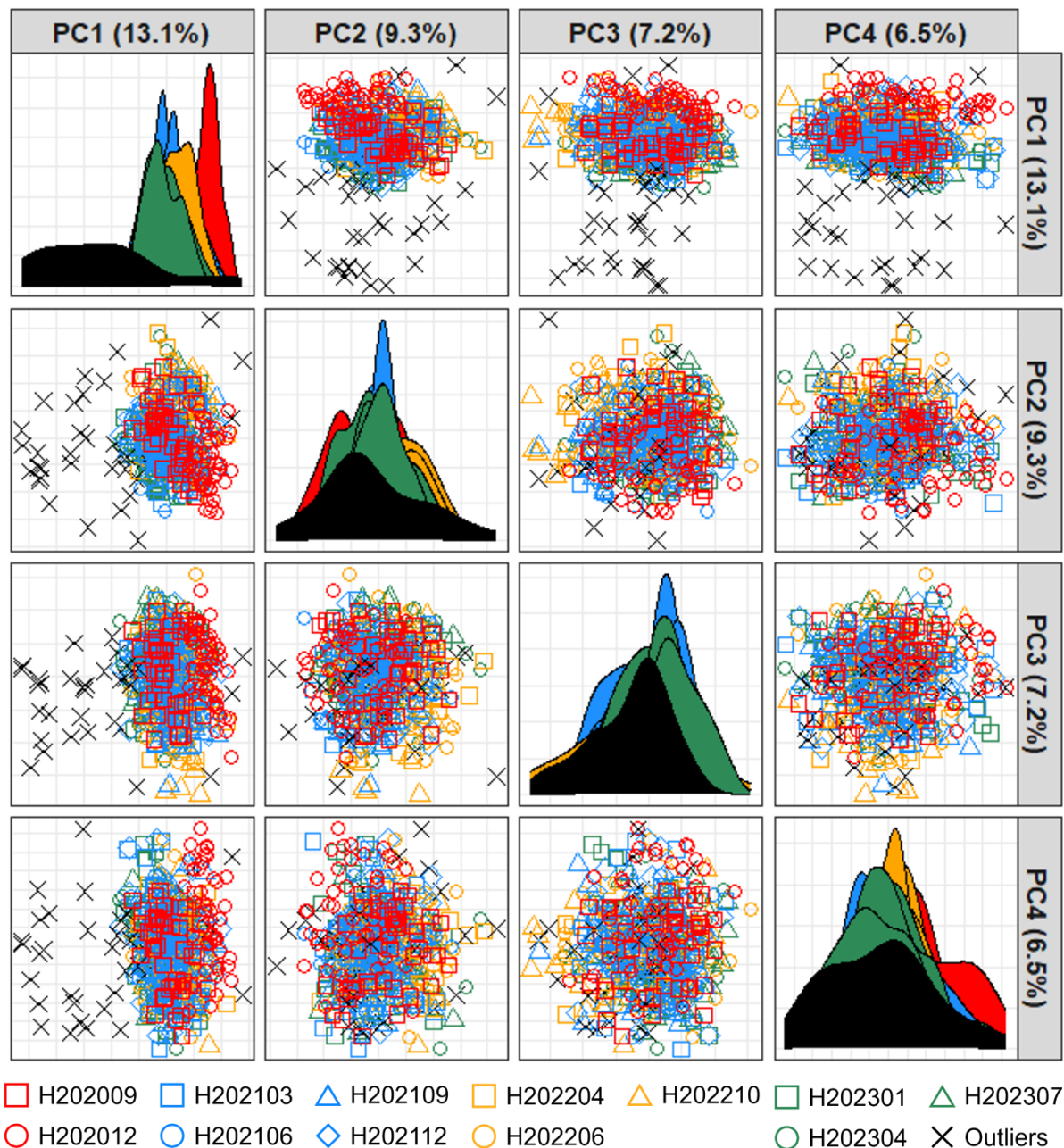
Figure 2 shows pair plots of PC scores projected by clustering using the DBSCAN algorithm on step and angle data with outliers. The DBSCAN algorithm detected 34 outlier data points. As shown in the pair plots, the detected data points were not grouped into any clusters and were located far away. The detected outliers comprised four from H202009, six from H202009, one from H202106, three from H202112, three from H202204, two from H202206, eight from H202210, one from H202301, two from H202304, and four from M202307. The detected outliers were removed from the dataset. Consequently, the dataset became an imbalanced dataset. The effect of outliers on classification performance was evaluated in the classification modeling section.

Figure 3 shows the PC score plots of the first two PCs derived from the step and angle data with and without outliers. In Fig. 3a, most data points form a large, unified cluster, with the exception of some data points in the PC1 low region. These data points are detected as outliers in Fig. 2. Outliers are not helpful relative to model construction because they add complexity and introduce confusion to the training process of machine learning models (Hwang *et al.* 2023; Lee *et al.* 2024a, b).

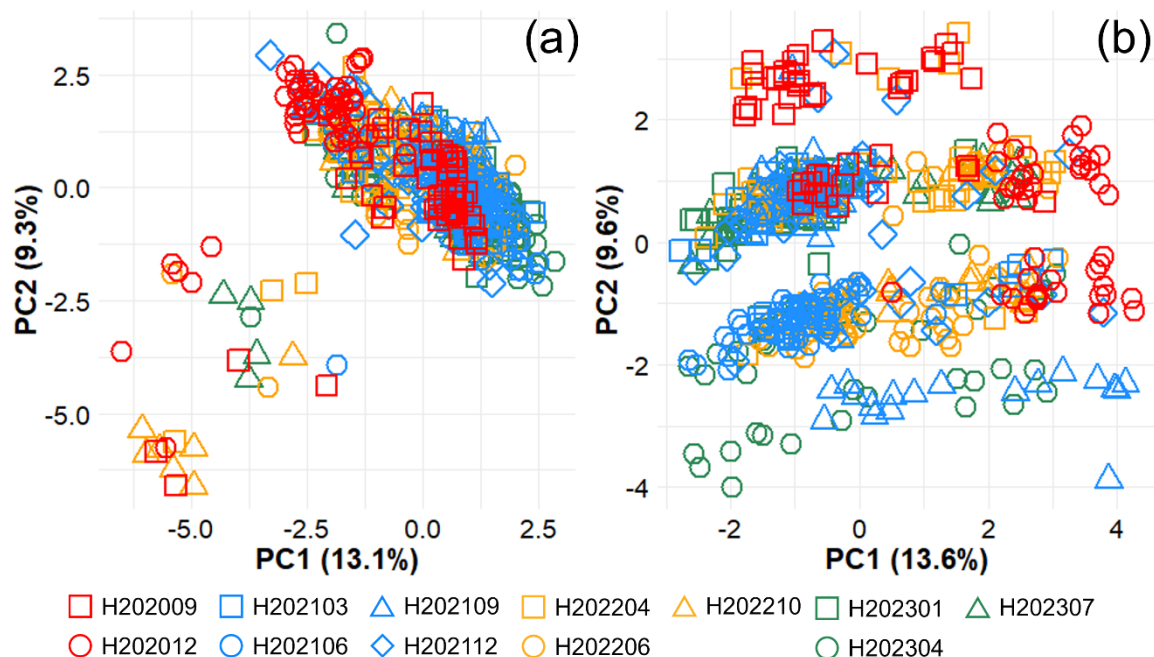
In preparation of Fig. 3b, the outliers had been removed from the dataset prior to PCA. The characteristics of the forming clusters changed due to outlier removal. The data points were separated from the large cluster shown in Fig. 3a and grouped into several distinct clusters. However, some data points were divided into two groups even though the product was manufactured on the same date. This indicates that the paper was manufactured using twin-wire systems, such as hybrid or gap formers. In the Fourdrinier wire system, the headbox slice jet impinges on a single forming fabric. Thus, the angle in the same step is observed like a mirror image on them. Such information provides evidence that the paper was probably manufactured using a Fourdrinier wire system. In contrast, in a twin-wire system, the headbox slice jet impinges into the converging gap between the two forming fabrics (Jeong *et al.* 2024a,b; Smook 2002). Consequently, different weave marks appear on both the top and bottom surfaces of the paper.

In forensic document examination, documents may contain content written on either the top or bottom side. Notably, for papers manufactured using twin-wire systems, such as hybrid or gap formers, it becomes challenging to distinguish between the wire and felt sides. This suggests that the dataset should be composed of separate classes for the top

and bottom sides. To address this issue, when constructing a dataset with step and angle data using a 2D lab formation sensor, it is essential to determine whether the patterns of the weave marks on the top and bottom sides are consistent.



**Fig. 2.** Pair plots of PC scores projected by clustering using DBSCAN on step and angle data with outliers



**Fig. 3.** PC score plots depicting the first two PCs derived from the step and angle data with outliers (a) and without outliers (b).

### Classification Models for Printing Paper Dating

It is well documented that, when analyzing periodic marks using a 2D-F sensor, the step and angle values corresponding to the top five intensity levels are particularly significant (Jeong *et al.* 2024a,b; Lee *et al.* 2024c). In this study, machine learning models were constructed using two types of input data: step and angle values from all intensity levels (unselected), and those from the top five intensity levels only (selected).

The classification performance and optimal hyperparameters of the three tree-based models are summarized in Table 4. For models trained with data derived from periodic marks, the angle values were preprocessed using the modulo 180 transformation described in Eq. 1. In all cases, models trained with selected variables outperformed those trained with unselected variables. These findings are consistent with previous reports (Jeong *et al.* 2024a,b; Lee *et al.* 2024c). The DT model exhibited minimal improvement with variable selection (F1 score of 0.627 to 0.629). The XGB model exhibited the best improvement with variable selection, increasing its F1 score from 0.657 to 0.747, whereas the hyperparameters remained consistent.

**Table 4.** Model Comparison for Printing Paper Dating

| Models | Variables               | Hyperparameters               | F1 score |
|--------|-------------------------|-------------------------------|----------|
| DT     | Unselected <sup>a</sup> | cp = 0.001                    | 0.627    |
|        | Selected <sup>b</sup>   | cp = 0.006                    | 0.629    |
| XGB    | Unselected <sup>a</sup> | lr = 0.1, max_depth = 7       | 0.657    |
|        | Selected <sup>b</sup>   | lr = 0.1, max_depth = 7       | 0.747    |
| RF     | Unselected <sup>a</sup> | n_feature = 1/3, n_trees = 91 | 0.761    |
|        | Selected <sup>b</sup>   | n_feature = 1/3, n_trees = 39 | 0.788    |

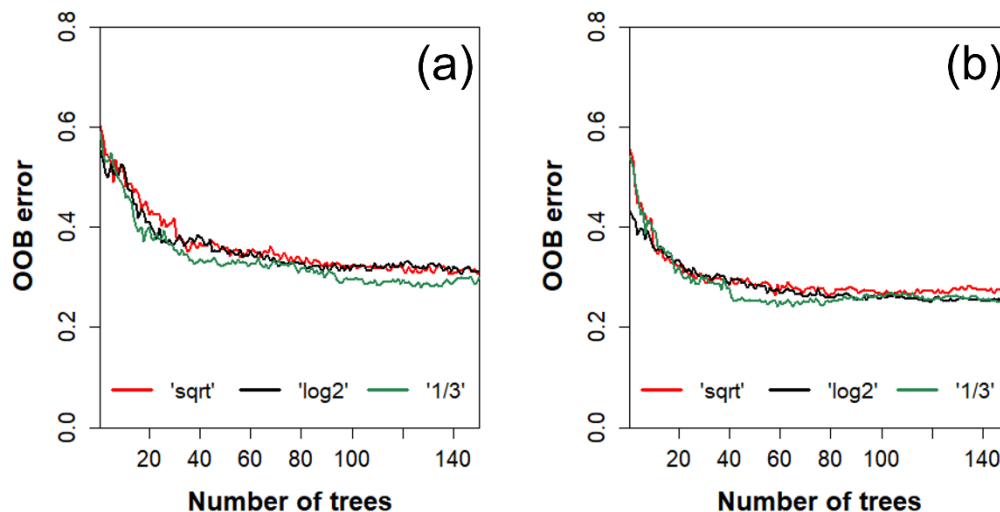
Definitions: a, step and angle data at all intensities; b, step and angle data at the top five intensities; cp: complexity parameter; lr: learning rate; max\_depth: maximum depth of the trees

The RF models outperformed the other classifiers, achieving F1 scores of 0.761 with unselected variables and 0.788 with selected variables. Overall, the models trained on the step and angle data at the top five intensities exhibited increased F1 scores. Thus, when analyzing periodic marks using a 2D lab formation sensor, the step and angle data at the top five intensities are significant. Among the tree-based algorithms, RF, which utilizes an ensemble learning process, was the most effective model. Thus, the RF model, which achieved the highest F1 score among the three classifiers, was selected for feature importance analysis.

### Feature Importance Measures from RF

Figure 4 shows the progression of OOB errors as each classification tree is incorporated during the RF training process for printing paper dating. Figures 4(a) and 4(b) present similar patterns in OOB error reduction. Initially, as more trees are added, the OOB errors decrease; notably, the use of step and angle data at the top five intensities accelerates this reduction during the early stages of tree addition. Subsequently, the errors stabilize at a consistent level once the number of trees exceeds 100 in Fig. 4a and 50 in Fig. 4b.

Based on these results, the optimal hyperparameters for the RF models were determined (Table 4), and the best-tuned models were also used for feature importance analysis. Notably, the minimum OOB errors achieved during training and the overall classification performance were higher when using the selected step and angle data from the top five intensities compared to using all intensity levels (Table 4).

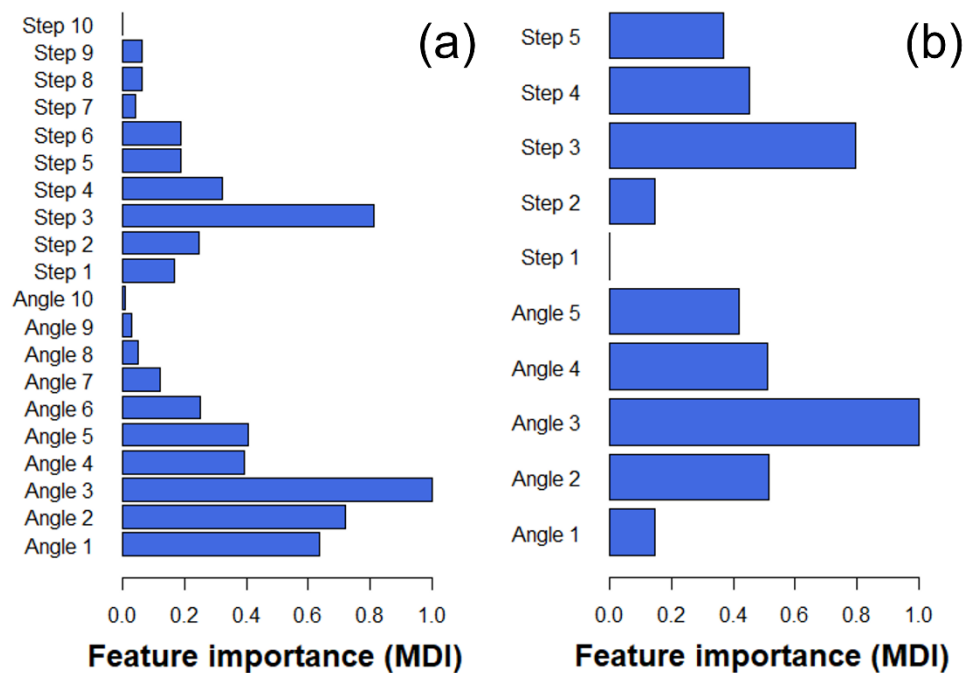


**Fig. 4.** Changes in out-of-bag (OOB) error rates with increasing number of classification trees in the dating of printing paper. Models trained with step and angle data at all intensity levels (a); Models trained with step and angle data at the top five intensity levels (b).

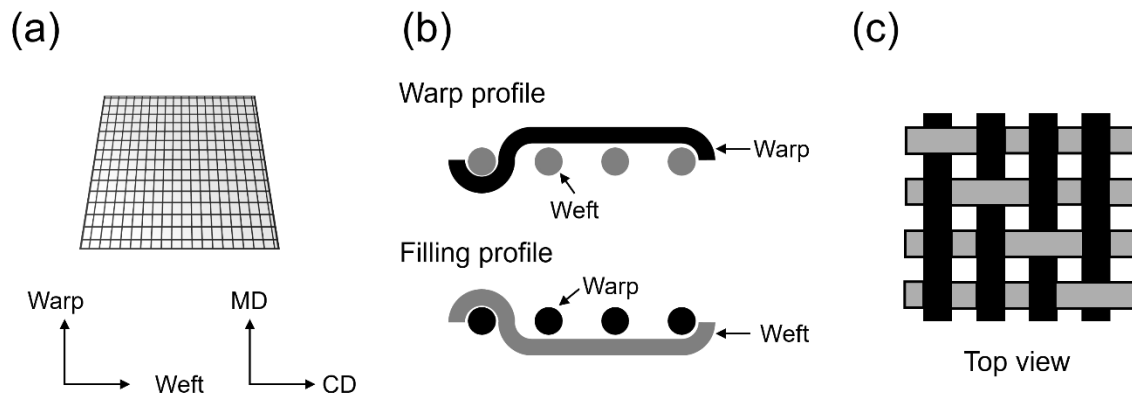
Figure 5 shows the feature importance of RF models trained on step and angle data. Figure 5a shows the feature importance of the step and angle at all intensities. It is clearly confirmed that the step and angle data at the third intensity level identified as most significant variable for printing paper dating. In general, the step and angle data at the first two intensities were measured for the weaving patterns of flat weaving (Jeong *et al.* 2024a, b; Lee *et al.* 2024c). Figure 6 illustrates the design of a simple forming fabric. The warp direction on the weaving machine aligns with the machine direction on the paper machine, whereas the filling (weft) direction on the weaving machine corresponds to the cross-machine direction (CD) on the paper machine (Adanur 2017). Therefore, angles near 360°



and 180° (Table 3) are general features of the weft direction (CD), which are observed for all types of products regardless of the manufacturing date.



**Fig. 5.** Feature importance of RF models trained on step and angle data at all intensities (a) and the top five intensities (b)



**Fig. 6.** Illustration of flat fabric weaving (a), profile of fabric design (b), and top view of forming fabric (c) (inspired (Adanur 2017))

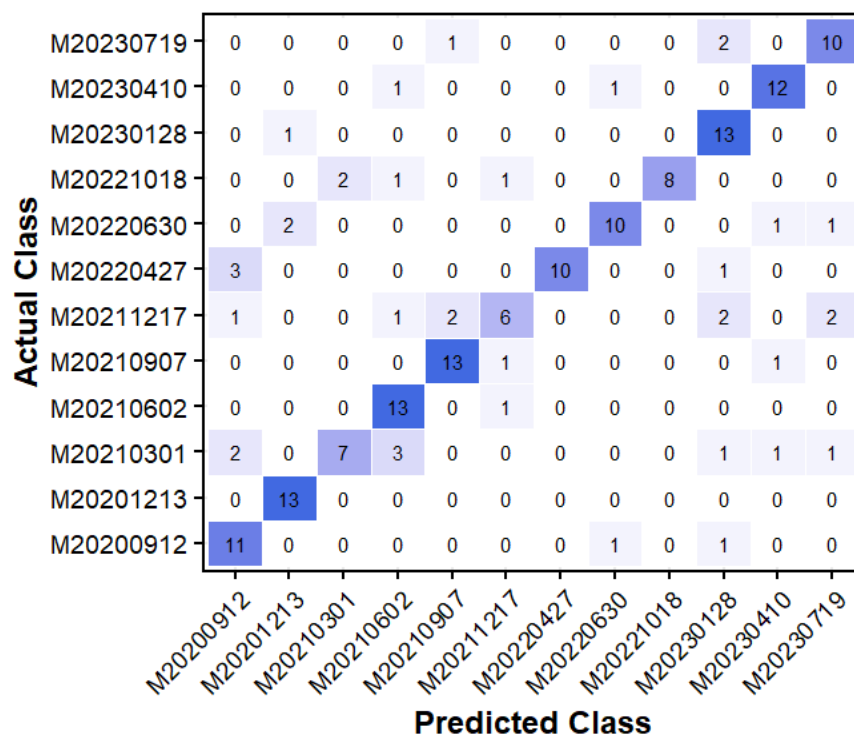
However, the step and angle data at the third intensity level (Table 3) were not associated with the weaving patterns. These data are assumed to represent the signatures of the outer warp knuckle (Adanur 2017) or drainage marks (Jeong *et al.* 2024a,b; Lee *et al.* 2024c). These characteristics depend on the type of forming fabric and the operation of the paper machine. The remaining variables likely exhibited low importance (Fig. 5), because they were either commonly observed across all paper samples, such as warp patterns, or they exhibited indistinct patterns due to lower intensities. Therefore, the step and angle data at the third intensity level were identified as the most significant features for forensic document examination using periodic marks on the paper. Based on these results, it can be hypothesized that if representative values—such as the step and angle data

at the third intensity level—match across samples with different production dates, it may indicate that the forming fabric was not changed during that period. Furthermore, a higher number of matching representative values between an unknown sample and a specific manufacturer's class increases the likelihood of identifying the manufacturer or approximating the production date of the sample.

### Misclassified Printing Paper

The final RF models trained on the step and angle data at the top five intensities exhibit low errors and high F1 scores (Table 4). These findings corresponded to those of previous studies (Jeong *et al.* 2024a,b; Lee *et al.* 2024c). However, the F1 score was very low, indicating that the model trained without intensity data performed poorly in printing paper dating. This limitation can be attributed to the structural characteristics of papers manufactured using a twin-wire system, as shown in the PCA results (Fig. 3b), where the periodic marks on the top and bottom sides are distinct. These findings indicate the need to treat the top and bottom sides as separate classes within the dataset. Therefore, when constructing a dataset using step and angle information obtained from a two-dimensional laboratory formation sensor, it is crucial to assess the consistency of weave mark patterns between the top and bottom surfaces.

Figure 7 shows the confusion matrix of the random forest (RF) model trained using step and angle data at the top five intensity levels. Based on the confusion matrix, the most frequently misclassified classes were H202103, H202112, and H202210. It is assumed that periodic marks from the top and bottom sides were mixed during dataset construction. In a twin-wire system, if different forming fabrics are used for the top and bottom wires, distinct periodic mark patterns can be imprinted on each side of the paper surface.



**Fig. 7.** Confusion matrix of the RF model trained using step and angle data at the top five intensity levels. Each cell represents the number of samples predicted for each class, with darker shades indicating higher frequencies.

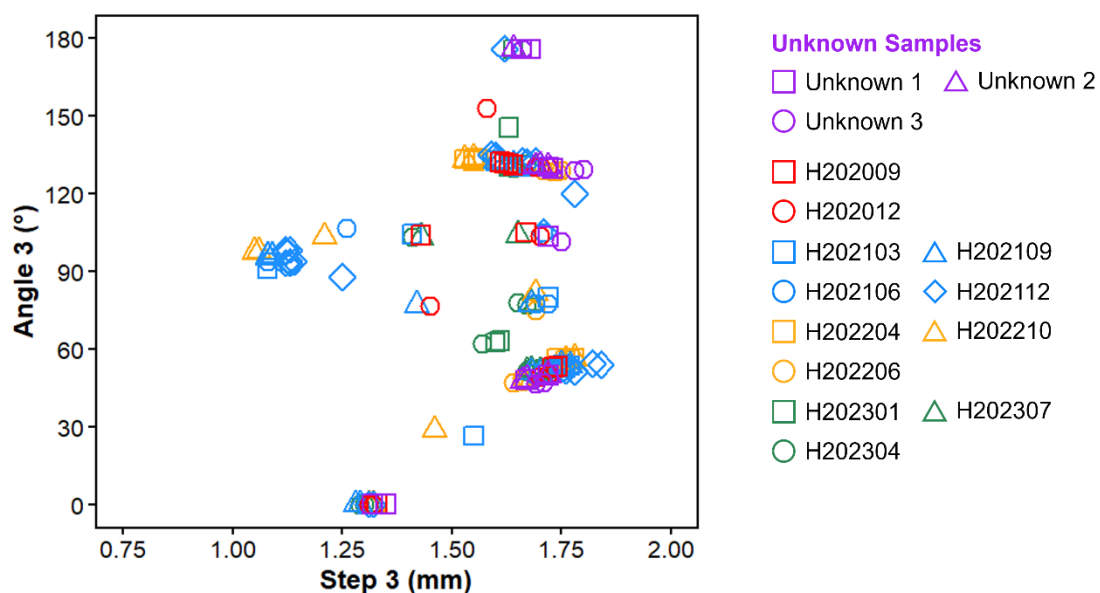


### Simplified Univariate Strategy for Printing Paper Dating

A total of 50 sheets were randomly sampled from office paper produced by the same manufacturer as the dataset, but with different production dates. The step and angle values at the third intensity peak were measured on both the top and bottom sides. Among the three unknown samples, Unknown 1 was produced in 2022 and Unknown 2 in 2020, whereas Unknown 3 showed no matches with any samples in the existing dataset. Additionally, it is assumed that the data points excluded from all clusters may have been influenced by fluctuations in intensity values, as such variations can alter the ranking of step and angle features based on intensity.

These results suggest that if a comprehensive dataset is available, a simple univariate approach using representative values such as the step and angle at the third intensity can be effectively applied for estimating the production period of printing paper. Furthermore, division into two subgroups within the same class was also observed using this method. This division may be attributed to the appearance of different weave marks on the top and bottom surfaces of the paper, likely due to the use of a twin-wire paper-making system.

From another perspective, when representative class values match between samples with different production dates, it can be inferred that the same forming fabric was used during that period. Therefore, continuously updating the representative value database over time can help identify the period during which forming fabric changes occurred. This approach can be further applied to estimate the date of production in the context of questioned document examination.



**Fig. 8.** Mapping unknown samples with known manufacturer but different production dates using step and angle values at the third intensity.

### Limitations and Future Studies

This study considered the dating of printing paper using periodic marks on its paper surface for forensic purposes. The dataset comprised the steps and angles of periodic marks on the paper surface. During dataset construction, the step and angle were measured randomly from either the top or bottom side of the paper, which may introduce bias into

the classification models. The periodic marks on the top and bottom sides of the paper differ if manufactured in a twin-wire system. In addition, if the manufacturer does not change the forming fabric used in production, it may not provide clear evidence for document dating. It is also common for machine operators to either regularly purchase the same brand and model of forming fabric or temporarily switch to a different brand before returning to the original. In such cases, periodic mark analysis can only reveal whether the same forming fabric was used, rather than providing a precise production date. On the other hand, if the repeated marks also exhibit features characteristic of felts, suction rolls, or dryer fabrics, the term “fingerprint” could be more robustly justified in forensic assessments of common origins of paper sheets.

These limitations should be addressed using sampling techniques and by increasing the data volume. When constructing a database, it is essential to determine whether the periodic marks on both sides are consistent. If differences exist, the dataset should be separated into top and bottom sides. Future work will focus on constructing a separated dataset to reduce bias and ensure that the model does not incorrectly recognize the periodic marks on the top and bottom sides as the same class.

## CONCLUSIONS

This study demonstrated that periodic marks on printing paper, particularly the step and angle data from light-transmitted images, can be effectively used for forensic document dating. The use of step-angle features from the top five intensity levels significantly improved classification performance in tree-based models. Among them, the random forest model achieved the highest F1 score. Feature importance analysis identified the third intensity level as the most critical for distinguishing manufacturing dates, likely reflecting forming fabric differences or machine-specific drainage marks. A simplified univariate strategy using these features further supported their forensic value. However, differences in periodic marks between the top and bottom sides of paper—common in twin-wire systems—may reduce model reliability. Therefore, future datasets should distinguish between paper surfaces to reduce bias and enhance interpretability. This study provides a foundation for the development of user-friendly and non-destructive tools for forensic document examination.

## ACKNOWLEDGMENTS

The authors would like to acknowledge the financial support provided by the Ministry of Science and ICT (Information and Communication Technology) of the Korean government and the National Research Foundation of Korea (Grant No. RS-2023-00301889).

## REFERENCES CITED

- Adanur, S. (2017). *Paper Machine Clothing: Key to the Paper Making Process*, CRC Press, Boca Raton, FL, USA. DOI: 10.1201/9780203744499
- Berger, C. E., and Ramos, D. (2012). “Objective paper structure comparison: Assessing

- comparison algorithms,” *Forensic Science International* 222(1-3), 360-367. DOI: 10.1016/j.forsciint.2012.07.018
- Berger, C. E. J. (2009). “Objective paper structure comparison through processing of transmitted light images,” *Forensic Science International* 192(1-3), 1-6. DOI: 10.1016/j.forsciint.2009.07.004
- Breiman, L. (2001). “Random forests,” *Machine Learning* 45(1), 5-32.
- Breiman, L. (2017). *Classification and Regression Trees*, Routledge, Abingdon-on-Thames, UK.
- Campello, R. J., Moulavi, D., and Sander, J. (2013). “Density-based clustering based on hierarchical density estimates,” in: *Pacific-Asia Conference on Knowledge Discovery and Data Mining*, Springer, Berlin, Heidelberg. DOI: 10.1007/978-3-642-37456-2\_14
- Chen, T., and Guestrin, C. (2016). “Xgboost: A scalable tree boosting system,” in: *Proceedings of the 22<sup>nd</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, CA, USA. DOI: 10.1145/2939672.2939785
- Choi, K. H., Lee, J. H., and Ryu, J. Y. (2018). “Analysis of modern document paper for identification of counterfeit document (I): Filler composition,” *Journal of Korea TAPPI* 50(3), 28-35. DOI: 10.7584/JKTAPPI.2018.06.50.3.28
- Comte, O., Dessimoz, D., Lanzi, L., Marquet, S., and Mazzella, W. J. P. (2006). “Paper discrimination by fast Fourier transform,” in: *4<sup>th</sup> European Academy of Forensic Science*, Helsinki, Finland.
- Ebara, H., Kondo, A., and Nishida, S. J. (1982). “Analysis of coated and non-coated papers by pyrolysis gas-chromatography,” *Rep. Natl. Res. Inst. Police Sci.* 2(35), 88-98.
- Ellen, D., Day, S., and Davies, C. (2018). *Scientific Examination of Documents: Methods and Techniques*, CRC Press, Boca Raton, FL, USA. DOI: 10.4324/9780429491917
- Ester, M., Kriegel, H.-P., Sander, J., and Xu, X. (1996). “A density-based algorithm for discovering clusters in large spatial databases with noise,” in: *KDD*, München, Germany. DOI: 10.5555/3001460.3001507
- Foner, H. A., and Adan, N. (1983). “The characterization of papers by X-ray diffraction (XRD): Measurement of cellulose crystallinity and determination of mineral composition,” *J. Forensic Sci. Soc.* 23(4), 313-321. DOI: 10.1016/S0015-7368(83)72269-3
- Grant, J. J. (1973). “The role of paper in questioned document work,” *J. Forensic Sci. Soc.* 13(2), 91-95.
- Havlíková, B., Katuščák, S., Petrovičová, M., Maková, A., and Brezová, V. (2009). “A study of mechanical properties of papers exposed to various methods of accelerated ageing. Part I. The effect of heat and humidity on original wood-pulp papers,” *Journal of Cultural Heritage* 10(2), 222-231. DOI: 10.1016/j.culher.2008.07.009
- Hwang, S.-W., Chung, H., Lee, T., Kwak, H. W., Choi, I.-G., and Yeo, H. (2023). “Investigation of NIR spectroscopy and electrical resistance-based approaches for moisture determination of logging residues and sweet sorghum,” *BioResources* 18(1), 2064-2082. DOI: 10.15376/biores.18.1.2064-2082
- Hwang, S.-W., Park, G., Kim, J., Kang, K.-H., and Lee, W.-H. (2024). “One-dimensional convolutional neural networks with infrared spectroscopy for classifying the origin of printing paper,” *BioResources* 19(1), 1633-1651. DOI: 10.15376/biores.19.1.1633-1651
- Jeong, C. W., Lee, Y. J., Chang Shin, Y., Choi, M. J., and Kim, H. J. (2024a). “Paper fingerprint by forming fabric: Analysis of periodic marks with 2D lab formation sensor and machine learning for forensic paper-identification,” *Nordic Pulp and*

- Paper Research Journal* 39(4), 747-757. DOI: 10.1515/npprj-2024-0050
- Jeong, C. W., Lee, Y. J., and Kim, H. J. (2024b). "Analysis of periodic marks for forensic document examination using boosting algorithms," *J. Korea Technical Assoc. Pulp and Paper Industry* 56(4), 65-73. DOI: 10.7584/JKTAPPI.2024.8.56.4.65
- Kuptsov, A. H. (1994). "Applications of Fourier transform Raman spectroscopy in forensic science," *J. Forensic Sci.* 39(2), 305-318. DOI: 10.1520/JFS13604J
- Lee, Y. J., Lee, T.-J., and Kim, H. J. (2024a). "Classification analysis of copy papers using infrared spectroscopy and machine learning modeling," *BioResources* 19(1). DOI: 10.15376/biores.19.1.160-182
- Lee, Y. J., Won, S. Y., Park, S. B., and Kim, H.-J. (2024b). "Chemometric approaches for discriminating manufacturers of Korean handmade paper using infrared spectroscopy," *Heritage Science* 12(1), 1-12. DOI: 10.1186/s40494-024-01460-6
- Lee, Y. J., Jeong, C. W., and Kim, H. J. (2024c). "Paper fingerprint by forming fabric: Analysis of periodic marks with 2D lab formation sensor and artificial neural network for forensic document dating," *BioResources* 19(4), 7591-7605. DOI: 10.15376/biores.19.4.7591-7605
- Lee, Y. J., Kweon, S. W., Jeong, C. W., and Kim, H. J. (2025a). "Evaluating the performance of machine learning and variable selection methods to identify document paper using infrared spectral data," *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* 327, article 125299. DOI: 10.1016/j.saa.2024.125299
- Lee, Y. J., Jeong, C. W., and Kim, H. J. (2025b). "Forensic feature extraction of document paper using periodic marks: PCA and t-SNE for manufacturer discrimination and document dating," *Forensic Science International* 367, article 112348. DOI: 10.1016/j.forsciint.2024.112348
- Louppe, G., Wehenkel, L., Sutera, A., and Geurts, P. (2013). "Understanding variable importances in forests of randomized trees," in: *Advances in Neural Information Processing Systems* 26, Lake Tahoe, Nevada, USA.
- Lundberg, S. M., and Lee, S. I. (2017). "A unified approach to interpreting model predictions," in: *Advances in Neural Information Processing Systems* 30, Long Beach, California, USA.
- Małachowska, E., Dubowik, M., Boruszewski, P., and Przybysz, P. (2021). "Accelerated ageing of paper: Effect of lignin content and humidity on tensile properties," *Heritage Science* 9, 1-8. DOI: 10.1186/s40494-021-00611-3
- Miyata, H., Shinozaki, M., Nakayama, T., and Enomae, T. (2002). "A discrimination method for paper by Fourier transform and cross correlation," *Journal of Forensic Sciences* 47(5), 1125-1132. DOI: 10.1520/JFS15491J
- Rhein, F., Hibbe, L., and Nirschl, H. (2024). "Hybrid modeling of hetero-agglomeration processes: A framework for model selection and arrangement," *Engineering with Computers* 40(1), 583-604. DOI: 10.1007/s00366-023-01809-8
- Smook, G. A. (2002). *Handbook for Pulp and Paper Technologists*, A. Wilde, Vancouver, Canada, 227-262.
- Spence, L. D., Francis, R. B., and Tinggi, U. J. (2002). "Comparison of the elemental composition of office document paper: Evidence in a homicide case," *Journal of Forensic Sciences* 47(3), 648-651.

Article submitted: December 8, 2024; Peer review completed: April 12, 2025; Revisions accepted: August 12, 2025; Published: August 28, 2025.  
DOI: 10.15376/biores.20.4.9208-9225