

Paper Fingerprint by Forming Fabric: Analysis of Periodic Marks with 2D Lab Formation Sensor and Artificial Neural Network for Forensic Document Dating

Yong-Ju Lee,^a Chang Woo Jeong,^b and Hyoung Jin Kim ^{a,*}

The increasing rates of illicit behaviors, particularly financial crimes, e.g., bank fraud and tax evasion, adversely affect national economies. In such cases, using nondestructive methods, scientists must evaluate relevant documents carefully to preserve their value as evidence. When forensic laboratories analyze paper as evidence, they typically investigate its origin and date of manufacture. If a document's date is earlier than the earliest availability of the paper used in its creation, then this anachronism indicates that the document has been backdated. This study investigated weave marks and drainage marks for forensic purposes. Machine learning models for forensic document examination were developed and evaluated. The partial least squares discriminant analysis (PLS-DA), support vector machine (SVM), and artificial neural network (ANN) classification models achieved F1-scores of 0.903, 0.952, and 0.931, respectively. In addition, to enhance model effectiveness and construct a robust model, variables were selected using the VIP scores generated by the PLS-DA model. As a result, the SoftMax classifier in the ANN model maintained its performance with an F1-score of 0.951 even with a 50% reduction in the number of input variables.

DOI: [10.15376/biores.19.4.7591-7605](https://doi.org/10.15376/biores.19.4.7591-7605)

Keywords: Forensic document dating; Copy paper; Classification; Artificial neural network (ANN); Support vector machine (SVM)

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 author: hyjikim@kookmin.ac.kr

INTRODUCTION

Criminals have access to various sophisticated technologies, and as science and technology improve, criminals become increasingly experienced at integrating technologies into their unlawful activities and motives to commit crimes. In the document examination field, which involves analyzing documents to differentiate between genuine and forged documents, identifying and authenticating paper documents, and determining their relative ages and dates, forensic document experts face an increasingly complex challenge as increasingly complicated criminal cases emerge. Increased criminal activity has an economic impact through tax evasion, bank fraud, counterfeiting, and other financial crimes. When this occurs, the questioned documents must be evaluated carefully to preserve their evidence value using nondestructive techniques. When a paper document is analyzed in a forensic laboratory, the most common processes involve determining when and where it was manufactured. Without eyewitnesses, physical or chemical evidence alone cannot determine the precise time a document was written accurately. This study is

based on an evaluation of the static properties of a document in terms of its alleged date. The date at which the paper was originally made available is compared to the document's date. The earliest availability of any materials used must be earlier than the date of the document. If the document's date is earlier than the earliest availability of the paper used in its creation, then such an anachronism is consistent with the document being backdated (Gupta 2018).

Traditionally, the paper discrimination task involves assessing various physical and optical properties, *e.g.*, tensile strength, thickness, basis weight, ash content, color, and fluorescence (Grant 1973). However, it is difficult for such methods to properly and dependably match two sheets of paper (Schlesinger and Settle 1971), thereby necessitating large sample sizes. As technologies have advanced, numerous paper analysis methods have been proposed, including X-ray diffraction (Foner and Adan 1983; Spence *et al.* 2000; Bisesi *et al.* 2006; Ellen *et al.* 2018;), elemental analysis (Spence *et al.* 2000; Spence *et al.* 2002), infrared spectroscopy (Andrasko 1996; Kher *et al.* 2001; Kher *et al.* 2005), Raman spectroscopy (Kuptsov 1994), image analysis (Miyata *et al.* 2002), and pyrolysis gas chromatography (Ebara *et al.* 1982).

In the papermaking process, pulp stocks are impinged onto forming fabric by a headbox slice jet. Pulp stocks include refined pulp, internal sizing agents, retention aids, fillers, defoamers, optical brighteners, and dry strength adhesives dissolved and suspended in aqueous solution (solid content: 0.5%). When the headbox jet (*i.e.*, a fiber suspension) impinges to drain the stock and form a fibrous web on the forming surface, an impression, *i.e.*, a weave mark or wire mark, is left on the sheet of paper by the mesh of the forming fabric. Drainage marks occur when the water drains, which leaves fibers on the forming fabric. Thus, these weave marks or drainage marks exhibit unique characteristics (similar to human fingerprints), depending on the brand of the forming fabric or changes in the wet-end process. These specific features can be utilized as forensic evidence for document examination. For example, Lee *et al.* (2023) investigated the potential of forensic feature extraction from forming fabric marks and formation using video spectral comparator (VSC) images. To identify VSC images based on paper document products, texture features extracted from images were converted using gray-level co-occurrence matrix methods, and a convolutional neural network model was tested. This method was used to classify seven major paper brands in the Korean market, and it achieved an accuracy of 97.66%. In addition, Berger (2009) developed a method to identify document paper using light transmission images and frequency analysis. A technical validation was carried out with 25 different papers, showing the potential of this method with common copy papers. This comprehensive analysis demonstrates that these specific features can be utilized effectively as forensic evidence for questioned document examination. However, such analyses require a sufficient database for comparison.

In this study, weave marks and drainage marks were investigated for forensic purposes. A literature review confirmed that it is possible to identify document paper based on the manufacturer; however, the potential for document dating has not been investigated extensively. In addition, for comprehensive forensic examination of document paper, a large database is required to perform comparative analyses. Machine learning models for forensic document examination have been developed and evaluated. This paper presents the results of a machine learning-based forensic identification method for document dating using paper fingerprints.

EXPERIMENTAL

Materials

Information about the copy paper utilized in this study is shown in Table 1. To date document paper, a total of 11 products with different production dates were collected from the same manufacturer.

Table 1. Information about the Copy Paper with Different Production Dates

NO.	Product	Code	Manufacture date (MM.YY)	Grammage (g/m ²)
1	"M" products from a Paper Company	M201311	11.2013	80
2		M201511	11.2015	
3		M201701	01.2017	
4		M201707	07.2017	
5		M201806	06.2018	
6		M201907	07.2019	
7		M202008	08.2020	
8		M202105	05.2021	
9		M202206	06.2022	
10		M202305	05.2023	
11		M202404	04.2024	

Dataset

Data formatting

A 2D-F sensor (Techpap, France) was used to scan both sides of each piece of copy paper. This data collection process was repeated to obtain 50 or 100 samples. The dataset is described in Table 2. The 2D-F sensor generates digital look-through images automatically, which are then analyzed using a fast Fourier transform algorithm to examine three distinct factors in the image, *i.e.*, intensity, angle, and step. Figure 1 shows the process used to calculate the intensity, angle, and step data from the obtained look-through images. Note that the angle and step data of the periodic marks were measured at the 10 highest intensities, resulting in 30 arrays in each measurement.

Table 2. Number of Papers Analyzed for Dataset Construction

NO.	Code	Number of Measured	
		Top side	Bottom side
1	M201311	100	100
2	M201511	100	100
3	M201701	100	100
4	M201707	100	100
5	M201806	100	100
6	M201907	100	100
7	M202008	100	100
8	M202105	50	50
9	M202206	50	50
10	M202305	50	50
11	M202404	50	50

The data scanned from each side of the paper were rendered as a 100 × 30 or 50 × 30 matrix. In this study, the sides of the paper are referred to as "Top" and "Bottom" to avoid assumptions about which side corresponds to the wire side or the felt side, as

traditionally identified in the papermaking process. This terminology is crucial because the exact orientation of the paper as it was manufactured is often unknown during forensic examination. Additionally, in cases where the paper was produced using a hybrid former, the distinction between the wire side and felt side may not be clear. Therefore, using “Top” and “Bottom” ensures clarity and consistency in the analysis.

Accordingly, the scanned matrices for the top and bottom of the paper were aligned horizontally. Figure 2 shows examples of how the experimental dataset was formatted. Ultimately, a dataset comprising 900 samples was constructed for classification modeling.

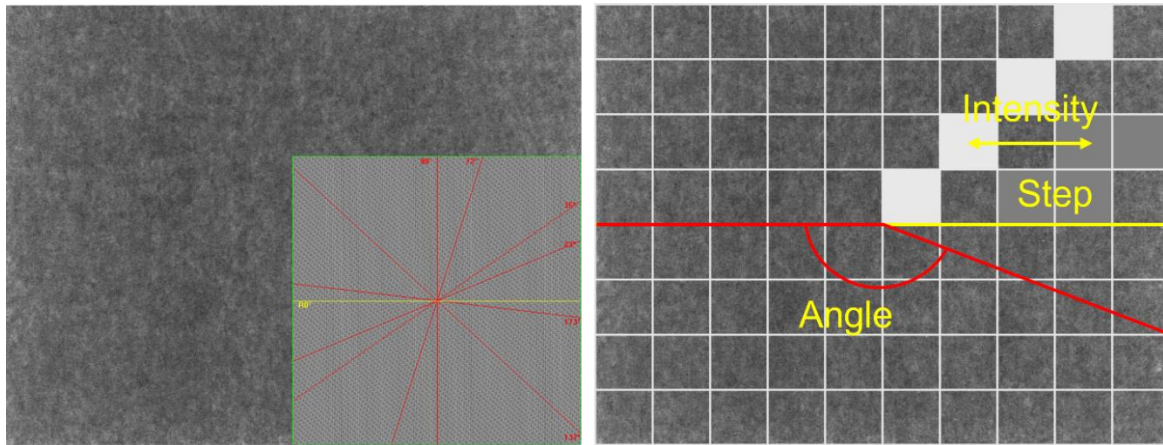


Fig. 1. Principles to calculate intensity, step, and angle measurements from digital look-through images

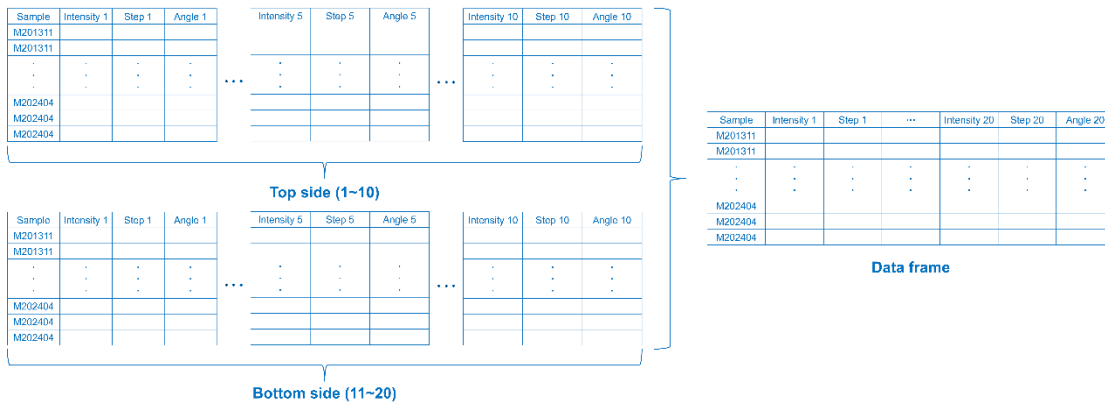


Fig. 2. Combining two data frames to format the dataset

Data preprocessing

It is possible to consider angles as symmetric (*e.g.*, 45° and 135°) within the same step, and in this case, they are considered equivalent. Thus, all angle data were converted using the absolute value of the sine function.

Dataset splitting

The dataset was split into training and test sets at a ratio of 7:3 to train and evaluate the compared classification models. Here, the data were partitioned using stratified random sampling to maintain the specified split ratio for all classes.

Classification Modeling

Partial least squares discriminant analysis

Partial least squares discriminant analysis (PLS-DA) is based on the PLS2 algorithm (Höskuldsson 1988; Wold *et al.* 2001; Stocchero 2019). PLS-DA is performed to identify latent variables that exhibit the maximum covariance between the independent variables (X) and their corresponding dependent variables (Y). This algorithm enables the extraction of relevant information from high-dimensional datasets while addressing potential multicollinearity among predictors. Here, the high-dimensional data were transformed into a new orthogonal coordinate system with 10 PLS-DA components, thereby effectively reducing the dimensionality of the data. Note that the optimization of PLS-DA models involves grid searches to control the number of PLS components, thereby minimizing the misclassification of the training data.

Support vector machine

The support vector machine (SVM) is designed to find a decision boundary that maximizes the margin between different classes, thereby facilitating linear classification. By utilizing a radial basis function (RBF) kernel, an SVM model projects data into a high-dimensional feature space to determine the optimal hyperplane (Vert *et al.* 2004). In this study, the cost and gamma parameters of the RBF kernel in the SVM were fine-tuned using grid searches over a logarithmic scale ranging from 2^{-3} to 2^3 for cost and 10^{-5} to 10^5 for gamma. The cost and gamma parameters manage the cost associated with misclassifying the training data and the Gaussian kernel used for nonlinear classification, respectively.

Artificial neural network

Artificial neural network (ANN) classifiers have been developed for document dating, and feedforward neural networks using the backpropagation algorithm have been utilized as classifiers (LeCun *et al.* 2012). A multilayer perceptron was implemented with the rectified linear unit as the activation function and cross-entropy as the loss function, and both the stochastic gradient descent and Adam optimizers were employed.

Table 3. ANN Architectures and Number of Nodes in each Layer

Shallow Network			Deep Network			
Input layer	Hidden layer	Output layer	Input layer	Hidden layer		Output layer
				1st	2nd	
60	16	11	60	16	16	11
	32			16	32	
	64			16	64	
	128			16	128	
				32	16	
				32	32	
				32	64	
				32	128	
				64	16	
				64	32	
				64	64	
				64	128	
				128	16	
				128	32	
	128	64				
	128	128				

To identify the optimal network architecture, various ANN configurations comprising either one or two hidden layers with different numbers of nodes were tested, and initial learning rates of 0.001, 0.01, and 0.1 were evaluated, with a maximum of 300 iterations for training. The network architecture, optimizer, and learning rate were fine-tuned through a grid search method utilizing threefold cross-validation. Then, the final classification model was selected based on the configuration that yielded the minimum loss. The architectures of the tested ANN model and the number of nodes in each layer are shown in Table 3.

Principal component analysis

Principal component analysis (PCA) was conducted to visualize the optimized model. Using PCA, the high-dimensional data were transformed into a new orthogonal coordinate system comprising five PCs. Then, the transformed data were visualized in a 2D space to analyze the patterns in the 2D-F sensor data.

Variable Importance Measures

The variable importance in projections (VIP) metric (Wold *et al.* 1993) from the PLS-DA model is commonly referred to as the VIP score (Ericksson *et al.* 2001). The concept behind this measure is to accumulate the importance of each variable j as reflected by w from each component. The VIP measure v_j is defined as follows.

$$v_j = \sqrt{\frac{p}{\sum_{a=1}^A SS_a} \sum_{a=1}^A (SS_a \cdot \left(\frac{w_{aj}}{\|w_a\|}\right)^2)} \quad (1)$$

where SS_a is the sum of the squares explained by the a^{th} component. Thus, the v_j weights measure the contribution of each variable according to the variance explained by each PLS component, where $\left(\frac{w_{aj}}{\|w_a\|}\right)^2$ represents the importance of the j th variable. The variance explained by each component can be computed by the expression $q_a^2 t_a' t_a$ (Ericksson *et al.* 2001); thus, v_j can also be expressed as follows.

$$v_j = \sqrt{\frac{p \cdot \sum_{a=1}^A q_a^2 t_a' t_a \cdot \left(\frac{w_{aj}}{\|w_a\|}\right)^2}{\sum_{a=1}^A SS_a}} \quad (2)$$

This provides a robust measure of the relative importance of each variable in the PLS model.

Evaluation Metrics

In classification tasks, it is essential to assess the classification accuracy into positive and negative categories. Here, true positives refer to correctly classified observations that belong to the positive class, and true negatives are correctly classified observations that belong to the negative class. In addition, false negatives are instances of positive classes incorrectly classified as negative, and false positives are instances of negative classes incorrectly classified as positive.

From these values, various performance indicators can be calculated to evaluate the classifier's ability to detect the target class (DeVries *et al.* 2003; Nielsen 2013). In this study, the F1-score was used to evaluate the classification performance of the models. In the classification of imbalanced datasets, the accuracy metric frequently yields biased results due to oversampled classes (Hwang *et al.* 2024). Thus, the F1-score, which is the harmonic mean of precision and recall, is more appropriate than accuracy. The precision, recall, and F1-score metrics are calculated as follows.

$$\text{Precision} = \frac{TP}{TP+FP} \quad (3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4)$$

$$F1 = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (5)$$

All data processing and classification modeling were performed using the R statistical software (R Core Team, ver. 4.4.1, Auckland, New Zealand).

RESULTS AND DISCUSSION

Angle and Step

Table 4 presents the top three intensity measurements, along with corresponding angle and step data, observed on the top side of the samples, while Table 5 provides the corresponding data for the bottom side. In forensic paper analysis, it is crucial to note that the terms "Top" and "Bottom" are used to describe the two surfaces of the paper without making assumptions about which side is the wire side or the felt side. This is because, during forensic examination, the exact orientation of the paper as it was produced is often unknown. Moreover, for papers produced by hybrid formers, the distinction between wire and felt sides may be further obscured, making the use of "Top" and "Bottom" even more appropriate.

Angles of 180°, which align with the cross direction (CD) of the papermaking process, are likely indicative of weave marks from the forming fabric. Similarly, vertical angles such as 90°, corresponding to the machine direction (MD), can also be attributed to weave marks. Moreover, specific angles like 125°, 53°, and various step sizes highlight distinct characteristics that vary depending on the manufacturer or papermaking process. These differing angles and step data reflect drainage marks, which function as unique identifiers of a manufacturer's papermaking machinery, akin to human fingerprints.

PLS-DA

The score plot (Fig. 3) depicts the first two PLSs derived from the intensity, step, and angle data.

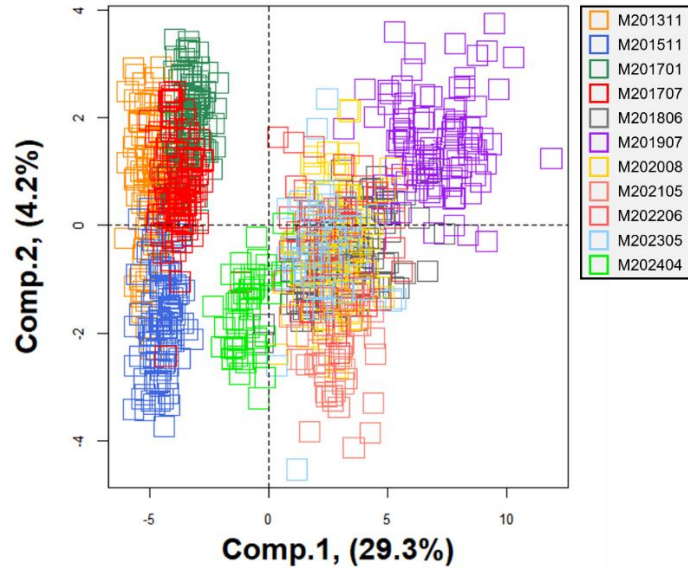


Fig. 3. PLS-DA score plot showing the first two PLSs derived from the intensity, step, and angle data

As can be seen, the first two components of PLS accounted for 29.3% and 4.2% of the covariance in the dataset, respectively. The PLS-DA score plots (Fig. 3) also show that while some samples were grouped into distinct clusters, which were made in 2013 to 2017, most of the samples' data points were mixed and formed a large, unified cluster.

The PLS-DA is a dimensionality reduction model that relies on linear combinations derived from various variables to classify each data class linearly. Consequently, it has certain limitations in terms of clarifying complex or multifaceted data structures. Thus, it is crucial to recognize the need to utilize more robust models, *e.g.*, SVM or ANN models, that can handle nonlinear classification.

Table 4. Intensity, Angle, and Step on Top Side

No.	Intensity 1	Angle 1, °	Step 1, mm	Intensity 2	Angle 2, °	Step 2, mm	Intensity 3	Angle 3, °	Step 3, mm
1	3.5	179.4	1.27	3.3	162.4	2.98	2.5	360.0	1.27
2	11.4	179.4	1.29	5.9	360.0	1.29	4.2	97.6	2.82
3	17.4	179.4	1.31	13.0	125.5	1.77	12.5	151.0	1.48
4	81.9	179.4	1.32	44.5	360.0	1.32	22.7	55.7	1.76
5	67.5	179.4	1.31	54.3	126.7	1.74	49.2	47.5	1.60
6	66.9	179.41	1.32	40.7	360.0	1.32	15.3	128.9	1.75
7	37.8	179.4	1.31	14.6	360.0	1.31	14.2	52.7	1.73
8	28.3	179.4	1.31	15.3	360.0	1.31	9.8	53.3	1.74
9	84.3	179.4	1.31	37.5	360.0	1.31	22.8	53.8	1.72
10	31.1	179.4	1.32	25.9	48.2	1.67	19.7	360.0	1.32
11	24.4	179.4	1.32	13.9	360.0	1.32	4.0	78.4	3.22

Notes: 1: M201311; 2: M201511; 3: M201701; 4: M201707; 5: M201806; 6: M201907; 7: M202008; 8: M202105; 9: M202206; 10: M202305; 11: M202404

Table 5. Intensity, Angle and Step on Bottom Side

No.	Intensity 1	Angle 1, °	Step 1, mm	Intensity 2	Angle 2, °	Step 2, mm	Intensity 3	Angle 3, °	Step 3, mm
1	4.8	179.4	1.27	4.5	162.9	3.14	3.7	360.0	1.27
2	27.6	179.4	1.28	11.3	360.0	1.28	3.4	17.2	2.91
3	22.2	179.4	1.31	11.6	125.5	1.77	8.3	46.5	1.57
4	54.9	179.4	1.29	17.5	360.0	1.29	14.5	49.2	1.64
5	60.3	179.4	1.31	39.9	100.2	1.70	39.5	73.6	1.66
6	60.3	179.4	1.32	35.7	360.0	1.32	14.7	48.2	1.67
7	60.4	179.4	1.32	36.7	128.9	1.75	30.6	48.2	1.67
8	35.4	179.4	1.31	14.3	131.4	1.63	10.4	360.0	1.31
9	46.7	179.4	1.32	25.2	360.0	1.32	14.1	132.0	1.56
10	32.4	179.4	1.32	22.1	360.0	1.32	15.4	48.2	1.67
11	17.7	179.4	1.33	10.6	360.0	1.32	6.0	77.7	1.67

Notes: 1: M201311; 2: M201511; 3: M201701; 4: M201707; 5: M201806; 6: M201907; 7: M202008; 8: M202105; 9: M202206; 10: M202305; 11: M202404

Model Comparison

The classification performance of the PLS-DA, SVM, and ANN models is compared in Table 6. As mentioned previously, the PLS-DA is a linear model, and the SVM and ANN methods are nonlinear models. These three models achieved F1-scores of 0.903, 0.952, and 0.931, respectively. Compared to the PLS-DA model, the SVM and ANN models demonstrated superior classification performance. For optimal classification on the experimental dataset, nonlinear models, *e.g.*, SVMs or ANNs, are required; however, these models have higher computational costs than PLS-DA models.

Table 6. Model Comparison for Dating Document Paper with All Variables

Model	Optimal Parameter	F1-score
PLS-DA	PLS components = 10	0.903
SVM	gamma = 0.01, cost = 2 ²	0.952
ANN	hl_size = (128, 64), lr = 0.0001, optimizer = adam	0.931

Note: hl_size: hidden layer sizes; lr: learning rate

VIP score

The analysis of the PLS-DA model made it possible to identify the relevant and crucial features for class differentiation. However, when handling a large number of initially detected features, there may be numerous sources of orthogonal noise or insignificant features that are irrelevant to the classification task. Thus, applying the classification model directly to the original data may lead to distortion because it could be affected by various types of noise and irrelevant features, thereby causing rotation in the original variable space. In such cases, the Pearson correlation coefficient ($p(\text{corr})$) values of each variable with the model components may also be influenced. As a result, some irrelevant features could appear to be relatively important for class separation, as determined by the VIP (Favilla *et al.* 2013; Galindo-Prieto *et al.* 2014). Thus, it is crucial to select informative features carefully when constructing a model to differentiate between different sample subpopulations (Xu *et al.* 2024).

The VIP scores for the top and bottom sides are shown in Figs. 4(a) and 4(b), respectively. Note that a high VIP value indicates that the variable plays a significant role in distinguishing between classes, and variables with low VIP values have a minimal impact on the model. As can be seen, for both sides, the first five values for intensity, step, and angle were identified as significant. These results underscore the effectiveness of the PLS-DA model. In contrast, the SVM and ANN models did not provide information about the decision-making process. Thus, the effectiveness of any model cannot be substituted directly.

Based on the results shown in Fig. 4, the first five values in the intensity, step, and angle data were identified as highly important features according to the VIP scores generated by the PLS-DA model. Note that reducing input variables is significant to construct a robust machine learning model. Excess input variables are a primary factor in increasing the model's computational cost (Hwang *et al.* 2024). The original datasets comprised 60 input variables, including 10 intensity, step, and angle data points for both the top and bottom sides. Based on the VIP scores (Fig. 4), the number of input variables was reduced to 30, including the first five values in the intensity, step, and angle data for both the top and bottom sides.

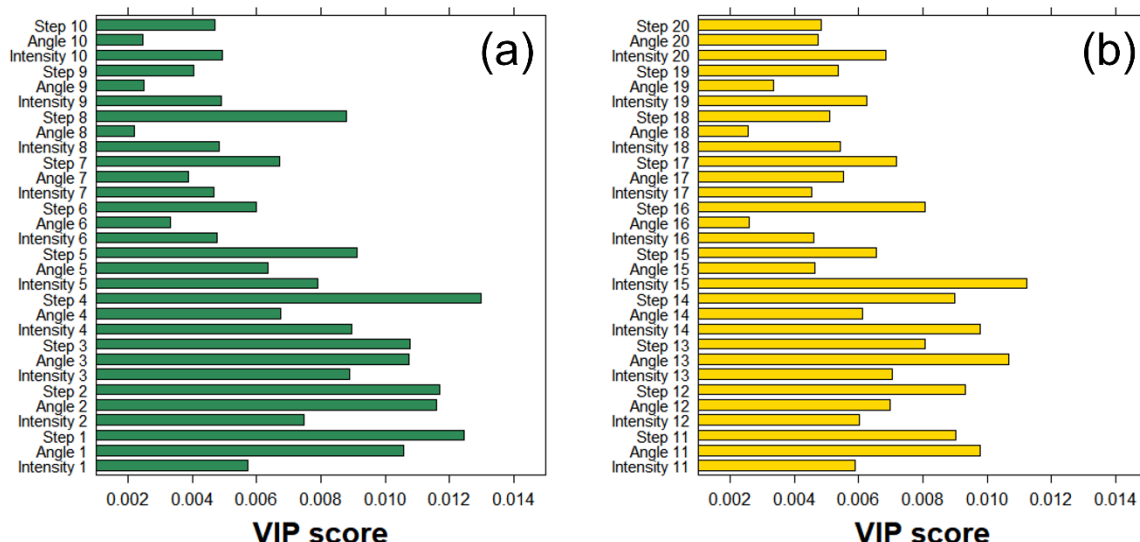


Fig. 4. VIP scores generated from PLS-DA on the (a) top side and (b) bottom side

Variable Selection

The classification performance (F1-score) of the three models trained with the selected variables is shown in Fig. 5. The reduction in the number of input variables led to lower classification performance for both the PLS-DA and SVM models, with F1-scores of 0.792 and 0.909, respectively. However, the SoftMax classifier in the ANN model yielded better results despite having a reduced number of input variables. These findings confirm the superior classification performance of the ANN model in the paper document dating task using the dataset obtained from the 2D-F sensor.

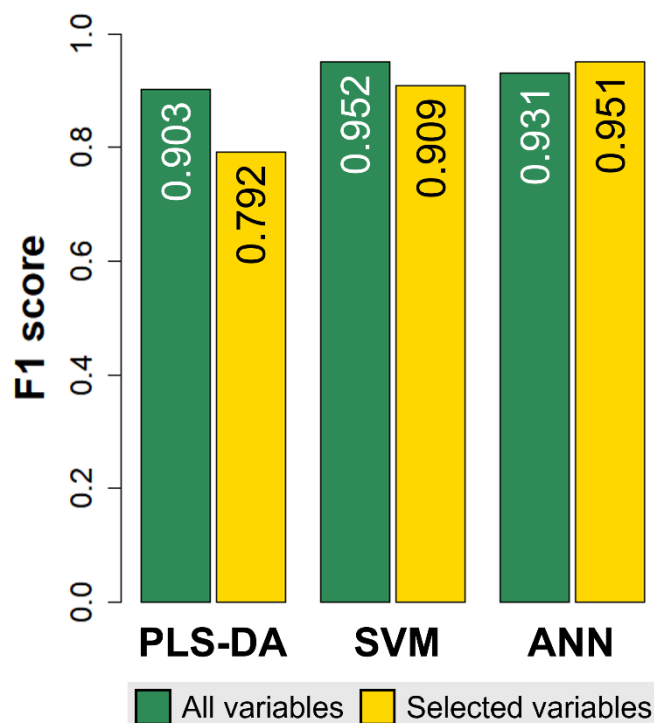


Fig. 5. Comparison of F1-scores based on variable selection

Dating of Unknown Documents

Unknown samples were scanned at 10 different points on the top and bottom of a piece of copy paper for machine learning data matching. In practice, this might be an effective

approach for forensic document examiners. In terms of forensic document inspection, once a dataset is formed, the date of production of the questioned document can be estimated. In addition, the time-consuming problem could be solved when the machine learning algorithm identifies a few candidates.

The ANN model was employed to predict the manufacture date of the same products, similar to forensic document examinations. Table 7 shows the predicted probabilities of the manufacture dates for three unknown products derived from the SoftMax classifier in the ANN model. Here, the model assigned a 93% probability to unknown product 1 of M202008 and predicted that unknown products 2 and 3 were of M202206 and M202105 with probabilities of 61% and 64%, respectively. These results demonstrate that identifying the manufacture date of questioned documents is possible if comprehensive databases are established.

This paper demonstrates how different types of paper can be identified by comparing the effectiveness of established models, assessing their performance in order to demonstrate the potential for distinguishing paper documents. However, there are limitations to collecting all the copy paper on a regular basis. Establishing a dataset by collecting copy paper once a month allows for the potential improvement of document dating accuracy through the use of deep learning algorithms with big data in forensic document examination.

Generally, the life of the forming fabric is one or two months, and new models with the latest technology are continuously developed and applied to improve the retention and dewatering ability. However, it would be impossible to infer when it was manufactured if a manufacturer did not change the model for a year or two. Furthermore, supplementing periodic mark analysis with additional information obtained from techniques such as infrared (IR) spectroscopy, Raman spectroscopy, ultraviolet-visible-near-infrared (UV-VIS-NIR) spectroscopy, colorimetry, and fluorescence spectroscopy could enhance the precision and overcome its limitations.

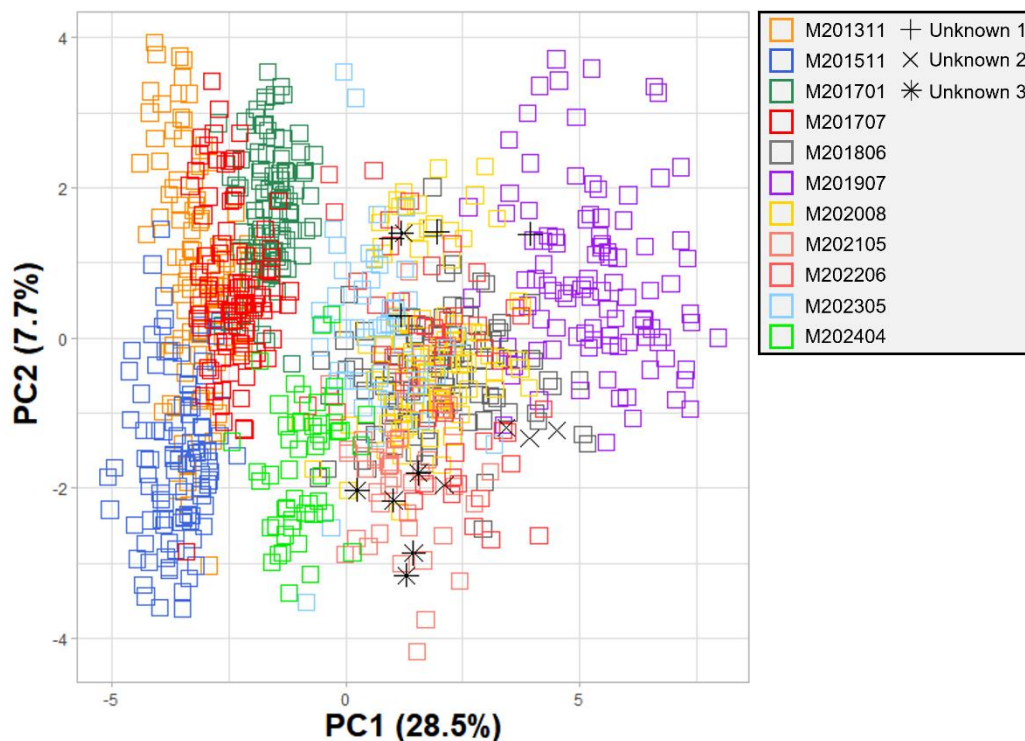


Fig. 6. PC score plot showing the initial two PCs obtained from the selected variables

Table 7. Predicted Probabilities for Document Dating of Unknown Samples using the ANN Model

Sample No.	1	2	3	4	5	6	7	8	9	10	11
Unknown 1	0.00	0.00	0.00	0.00	0.01	0.03	0.93	0.00	0.01	0.02	0.00
Unknown 2	0.00	0.00	0.00	0.00	0.12	0.14	0.11	0.01	0.61	0.00	0.00
Unknown 3	0.00	0.00	0.00	0.00	0.26	0.01	0.00	0.64	0.03	0.00	0.06

Notes: 1: M201311; 2: M201511; 3: M201701; 4: M201707; 5: M201806; 6: M201907; 7: M202008; 8: M202105; 9: M202206; 10: M202305; 11: M202404

Figure 6 shows the unknown products on the PC score plot with other products in the experimental dataset. As observed in Table 7, the data points of each unknown sample were assigned to the class groups on the PCA score plot that received strong support from the ANN classifier with high probability.

CONCLUSIONS

1. The collected data using the 2D-F sensor encompassed intensity, angle, and step information. The angles and steps imprinted on the surface of the paper during the papermaking process served as key indicators (*i.e.*, paper fingerprints) for dating the paper documents. These distinct features are largely determined by the time intervals for changing the consumable forming fabric.
2. The angle and step data of the periodic marks were measured at the 10 highest intensity levels, resulting in a 1×30 array for each measurement. These data were used to construct the classification model, with each dataset containing 60 variables representing the scanned matrices of both the top and bottom sides of the paper (aligned along the row direction). The data acquisition process was repeated until a total of 50 or 100 samples were collected. Ultimately, the classification model was constructed using data from a total of 900 samples
3. The PLS-DA, SVM, and ANN classification models trained on all 60 variables achieved F1-scores of 0.903, 0.952, and 0.931, respectively. To enhance effectiveness and construct a sufficiently robust model, variables were selected using the VIP scores generated by the PLS-DA model, corresponding to the initial five values of the intensity, step, and angle on both the top and bottom sides of the copy paper. The SoftMax classifier in the ANN maintained its performance with an F1-score of 0.951 even with a 50% reduction in the number of input variables.
4. The optimized ANN model was then employed to predict the production date of the document paper. The model assigned a 93% probability to unknown product 1 being from M202008 and predicted that unknown products 2 and 3 were from M202206 and M202105 with probabilities of 61% and 64%, respectively. These results demonstrate that identifying the manufacture date of questioned documents is feasible if sufficiently comprehensive datasets can be acquired.

ACKNOWLEDGMENTS

This study was carried out with the support of 'R&D Program for Forest Science Technology (Project No. RS-2024-00404816)' provided by Korea Forest Service (Korea Forestry Promotion Institute).

REFERENCES CITED

- Andrasko, J. (1996). "Microreflectance FTIR techniques applied to materials encountered in forensic examination of documents," *J. Forensic Sci.* 41(5), 812-823. DOI: 10.1520/JFS14003J
- Berger, C. E. H. (2009). "Objective paper structure comparison through processing of transmitted light images," *Forensic Sci. Int.* 192(1-3), 1-6. DOI: 10.1016/j.forsciint.2009.07.004
- Bisesi, M. S. (2006). "ASTM guidelines for forensic document examination," In: *Scientific Examination of Questioned Documents*, J. S. Kelly and B. S. Lindblom (eds.), CRC Press, Boca Raton, FL, USA, pp. 383-396.
- DeVries, T., Von Keyserlingk, M., Weary, D., and Beauchemin, K. A. (2003). "Validation of a system for monitoring feeding behavior of dairy cows," *J. Dairy Sci.* 86(11), 3571-3574. DOI: 10.3168/jds.S0022-0302(03)73962-9
- Ebara, H., Kondo, A., and Nishida, S. (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.
- Ericksson, L., Johansson, E., Kettaneh-Wold, N., and Wold, S. (2001). *Multi-and Megavariate Data Analysis: Principles and Applications*, Umetrics Academ, Umea, Sweden.
- Favilla, S., Durante, C., Vigni, M. L., and Cocchi, M. (2013). "Assessing feature relevance in NPLS models by VIP," *Chemom. Intell. Lab. Syst.* 129, 76-86.
- 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
- Galindo-Prieto, B., Eriksson, L., and Trygg, J. (2014). "Variable influence on projection (VIP) for orthogonal projections to latent structures (OPLS)," *J. Chemom.* 28(8), 623-632. DOI: 10.1002/cem.2627
- Grant, J. (1973). "The role of paper in questioned document work," *J. Forensic Sci. Soc.* 13(2), 91-95. DOI: 10.1016/s0015-7368(73)70774-x
- Gupta, R. R. (2018). "A scientific method for forensic examination of paper," *IP International Journal of Forensic Medicine and Toxicological Sciences* 3(2), 18-20. DOI: 10.18231/2456-9615.2018.0005
- Höskuldsson, A. (1988). "PLS regression methods," *J. Chemom.* 2(3), 211-228. DOI: 10.1002/cem.1180020306
- 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
- Kher, A., Mulholland, M., Reedy, B., and Maynard, P. (2001). "Classification of document papers by infrared spectroscopy and multivariate statistical techniques," *Appl. Spectrosc.* 55(9), 1192-1198. DOI: 10.1366/0003702011953199
- Kher, A., Stewart, S., and Mulholland, M. (2005). "Forensic classification of paper with infrared spectroscopy and principal components analysis," *J. Near Infrared Spec.* 13(4), 225-229. DOI: 10.1255/jnirs.540
- Kuptsov, A. H. (1994). "Applications of Fourier transform Raman spectroscopy in forensic science," *J. Forensic Sci.* 39(2), 305-318. DOI: 10.1520/JFS13604J
- LeCun, Y. A., Bottou, L., Orr, G. B., and Müller, K.-R. (2012). "Efficient backprop," in: *Neural Networks: Tricks of the Trade*, G. B. Orr, and K.-R. Müller (eds.), Springer, Berlin, Germany, pp. 9-50. DOI: 10.1007/3-540-49430-8_2

- Lee, J., Kim, H., Yook, S., and Kang, T. Y. (2023). "Identification of document paper using hybrid feature extraction," *J. Forensic Sci.* 68(5), 1808-1815. DOI: 10.1111/1556-4029.15330
- Miyata, H., Shinozaki, M., Nakayama, T., and Enomae, T. (2002). "A discrimination method for paper by Fourier transform and cross correlation," *J. Forensic Sci.* 47(5), 1125-1132. DOI: 10.1520/JFS15491J
- Nielsen, P. P. (2013). "Automatic registration of grazing behaviour in dairy cows using 3D activity loggers," *Appl. Anim. Behav. Sci.* 148(3-4), 179-184. DOI: 10.1016/j.applanim.2013.09.001
- Schlesinger, H., and Settle, D. M. (1971). "A large-scale study of paper by neutron activation analysis," *J. Forensic Sci.* 16(3), 309-330.
- Spence, L. D., Baker, A. T., and Byrne, J. P. (2000). "Characterization of document paper using elemental compositions determined by inductively coupled plasma mass spectrometry," *Chemom. Intell. Lab. Syst.* 15(7), 813-819. DOI: 10.1016/j.chemolab.2013.05.013
- Spence, L. D., Francis, R. B., and Tinggi, U. (2002). "Comparison of the elemental composition of office document paper: Evidence in a homicide case," *J. Forensic Sci.* 47(3), 648-651.
- Stocchero, M. (2019). "Iterative deflation algorithm, eigenvalue equations, and PLS2," *J. Chemom.* 33(10), article e3144. DOI: 10.1002/cem.3144
- Vert, J.-P., Tsuda, K., and Schölkopf, B. (2004). *A Primer on Kernel Methods*, MIT Press, Cambridge, MA, USA.
- Wold, S., Johansson, E., and Cocchi, M. (1993). "PLS: Partial least squares projections to latent structures," in: *3D QSAR in Drug Design: Theory, Methods and Applications*, Kluwer ESCOM Science Publisher, Dordrecht, Netherlands, pp. 523-550.
- Wold, S., Sjöström, M., and Eriksson, L. (2001). "PLS-regression: A basic tool of chemometrics," *Chemom. Intell. Lab. Syst.* 58(2), 109-130. DOI: 10.1016/S0169-7439(01)00155-1
- Xu, S., Bai, C., Chen, Y., Yu, L., Wu, W., and Hu, K. (2024). "Comparing univariate filtration preceding and succeeding PLS-DA analysis on the differential variables/metabolites identified from untargeted LC-MS metabolomics data," *Anal. Chim. Acta* 1287, article 342103. DOI: 10.1016/j.aca.2023.342103

Article submitted: July 23, 2024; Peer review completed: August 7, 2024; Revised version received: August 10, 2024; Accepted: August 11, 2024; Published: August 27, 2024.
DOI: 10.15376/biores.19.4.7591-7605