ORIGINAL ARTICLE

Real-time biomass feedstock particle quality detection using image analysis and machine vision

Chandrakanth Gudavalli[1] · Elizabeth Bose[2] · Bryon S. Donohoe[2] · David A. Sievers[1]

Abstract

A common and costly challenge in the nascent biorefinery industry is the consistent handling and conveyance of biomass feedstock materials, which can vary widely in their chemical, physical, and mechanical properties. Solutions to cope with varying feedstock qualities will be required, including advanced process controls to adjust equipment and reject feedstocks that do not meet a quality standard. In this work, we present and evaluate methods to autonomously assess corn stover feedstock quality in real time and provide data to process controls with low-cost camera hardware. We explore the use of neural networks to classify feedstocks based on actual processing behavior and pixel matrix feature parameterization to further assess particle attributes that may explain the variable processing behavior. We used the pretrained ResNet neural network coupled with a gated recurrent unit (GRU) time-series classifier trained on our image data, resulting in binary classification of feedstock anomalies with favorable performance. The textural aspects of the image data were statistically analyzed to determine if the textural features were predictive of operational disruptions. The significant textural features were angular second moment, prominence, mean height of surface profile, mean resultant vector, shade, skewness, variation of the polar facet orientation, and direction of azimuthal facets. Expansion of these models is recommended across a wider variety of labeled feedstock images of different qualities and species to develop a more robust tool that may be deployed using low-cost cameras within biorefineries.

Keywords Machine vision · Neural network · Biomass feedstock · Automated image analysis

1 Introduction

In the evolving bioenergy industry, pioneer lignocellulosic biomass processing plants have experienced difficulties achieving their design capacities and meeting their expected revenue generation [1]. One of the common challenges is the consistent handling and conveyance of feedstock materials, which can vary widely in their particle size, particle shape, chemical composition, and other physical and mechanical attributes. Specifically, the size and shape of material can play a major role in the flow properties that make gravity-fed and even force-fed hoppers and conveyors function inconsistently or fail altogether [2]. Universal equipment that works on more conventional materials such as cereal grains can experience bridging, jamming, and motor load spikes when trying to covey lignocellulosic feedstocks such as corn stover [3]. New equipment designs specific to the feedstock properties may be necessary to alleviate some of these challenges, and additional solutions to cope with varying feedstock qualities may include the application of advanced process controls to adjust equipment and even reject some feedstocks that do not meet a quality standard—both of which require online instrumentation for real-time measurement [4].

A tool that is portable, scalable, and cost-effective for the determination of biomass particle quality is fundamental to the implementation of advanced process controls and acceptance/rejection of feedstock at the processing plant gate and even

Chandrakanth Gudavalli and Elizabeth Bose contributed equally to this work.

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s13399-020-00904-w) contains supplementary material, which is available to authorized users.

✉ Bryon S. Donohoe
bryon.donohoe@nrel.gov

✉ David A. Sievers
david.sievers@nrel.gov

1 National Renewable Energy Laboratory: National Bioenergy Center, 15013 Denver West Parkway, Golden, CO 80401, USA

2 National Renewable Energy Laboratory: Bioenergy Sciences Center, 15013 Denver West Parkway, Golden, CO 80403, USA

Springer

grading feedstock before it is transported to the biorefinery. Online, real-time particle shape and sizing instruments are currently available and in use in several product processing industries [5, 6]. Although these solutions fill the need for online instrumentation reasonably well, cost may be a limiting factor in widespread adoption and at multiple locations in the processing line for automation, and limits to portability and hand-held use do not support in-field use. Another alternative to the high-quality, high-cost optical analysis option is the use of low-cost, bright-field imaging combined with image analysis and machine vision (MV) to assess feedstock variability. MV is not a new concept, but the last several years have seen a rapid increase in the development and maturation of this technology as computational power has evolved. Automated medical imaging analysis is an example of this growing area, where difficult-to-identify abnormalities can be detected more efficiently using models such as deep neural networks (NNs) to allow radiologists to better detect diseases [7, 8]. Simple instruments such as digital photographic cameras offer a similarly attractive option for assessing biomass particle quality in our current application. In this paradigm, low-cost cameras are placed over conveyor belts and hoppers and even employed as hand-held units such as smartphones to capture photographs of feedstocks, and a MV model assesses the images and output data used for intelligent decision-making in real time. Due to the low cost (< \$1000) and ubiquitous nature of digital cameras, this method may be easily scaled to many measurement points throughout a processing plant for the cost of a single installation of a formal commercial measurement hardware (e.g., laser scanners). The versatility of digital cameras means that additional features can be extracted from images, such as textural feature and color data, for assessing additional attributes. Finally, NN-based methods do not require absolute measurements and behavioral models to predict performance and make classifications (i.e., measured particle size distribution ported to a mechanistic model to determine if material will flow). Instead, they can be developed with end-use data to classify conditions (e.g., motor load spikes) to be optimized. These reasons encourage the pursuit of MV as a deployable online instrumentation technique for biorefinery operations.

In this work, we present and evaluate methods to autonomously assess corn stover feedstock quality in real time with low-cost camera hardware. We explore two different MV approaches here that include the use of neural networks to classify feedstocks based on actual processing behavior and pixel matrix feature parameterization (PMFP) to further assess particle attributes that may explain the processing behavior. These two approaches are complementary; as presented here, the NN-based approach has been applied to determine when feedstock is present that causes hopper motor load upsets, whereas the PMFP method offers a path to understand the underlying reasons for the upsets.

2 Methods

2.1 Image set and labeling

We operated a 500 kg/d-rated pilot plant at the National Renewable Energy Laboratory (NREL) Integrated Bioenergy Research Facility over four 50-h experimental runs to convert milled corn stover into fermentable sugars via a dilute-acid chemical hydrolysis and to observe equipment operability and performance [3]. Feeding equipment upstream (colored orange in Fig. 1a) experienced motor load (torque) upsets that were linked to varying feedstock quality fed from the conical feed-hopper. This unit operated continuously but was refilled approximately every 30 min by intermittent operation of upstream handling equipment. Equipment operators described incoming feedstock variability and particle size and morphology segregation within the conical feed-hopper between refill events, and these are hypothesized as causes for equipment load variations, which pose a risk to equipment reliability if not mitigated. Anomalous motor loads shown in Fig. 1b occurred in sync with conical feed-hopper refill events and lasted for several minutes each refill cycle.

Images of feedstock on a conveyor belt were captured by a GoPro HERO 6 camera modified with a rectilinear 5-mm f/1.4 lens set at f/8 at approximately 0.5-m focal distance above the weigh belt. Two 500-W halogen lamps provided illumination, and the camera automatically controlled exposure. Images were taken every 30 s, which resulted in approximately 26,000 total images (examples shown in Fig. 2). Weigh-belt edges were visible in the frames, with feedstock piled mostly in the center of the belt. Images were cropped, leaving only the center third of the frame where only feedstock was normally visible for further analysis. Despite dust control measures, most images taken during run 1 had a haze from lens dust accumulation, whose implications are discussed further below. Subsequent improvements sufficiently mitigated lens dust for the other three runs.

Samples of feedstock were only available in 1-h increments, posing a challenge for quantifying feedstock quality through particle size distribution, bulk density, or some other attribute, so images could not be effectively labeled with this information. Instead, actual equipment performance indicators recorded by the process automation system (supervisory control and data acquisition, or SCADA) were used to label the images. The cross-feeder (Fig. 1) is a simple live-bottom hopper that was operated at full speed as a conveyor. This unit provided a low-noise motor load signal that varied significantly with upstream conical feed-hopper segregation events. A rolling-window quantile threshold of 70% was used to distinguish between normal and anomalous operation, and using a calculated delay time between weight-belt and cross-feeder behavior, associated images were labeled based on timestamps as either TRUE (coincident with high cross-

Fig. 1 **a** Process flow diagram. Orange-colored units experienced variation in their performance caused by feedstock variability. Photographs of feedstock were captured every 30 s at the weigh belt. **b** Plot of cross-feeder torque during operation as example of anomaly labeling; times when torque exceeds the quantile value are highlighted in orange (CH, chemical hydrolysis; EH, enzymatic hydrolysis)
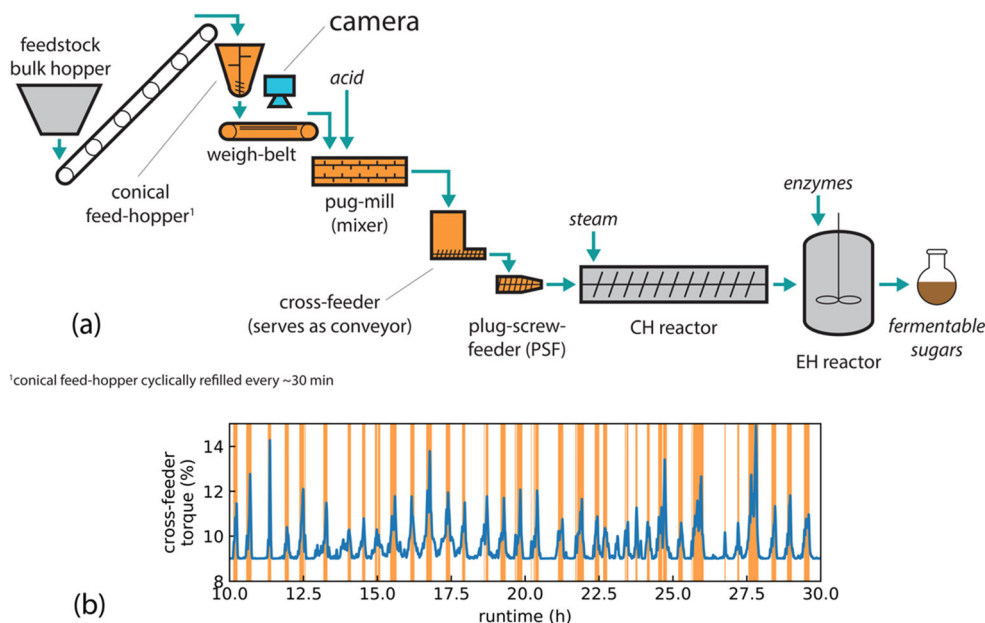
feeder torque) or FALSE (coincident with normal cross-feeder torque), as depicted in Fig. 1b. Table 1 summarizes the labeled images.

2.2 Neural network classification

Image classification remains a machine vision and image-processing challenge. Though there are classical image-processing techniques that can be used for classification, recent advancements in deep learning have resulted in several novel and efficient image classification models [9, 10]. With the advent of convolutional neural networks (CNNs), it has become possible to develop robust image classification and recognition models [11]. We present a CNN-based approach for image classification to detect anomalous feedstock while it is in view of a camera in process equipment, such as the weigh belt in our process. There are several standard CNN architectures like VGG-16, VGG-19, ResNet, Inception, LeNet, and Xception, which demonstrated good feature-recognition performance on standard image data sets like ImageNet, CIFAR-10, and COCO [12–14].

We generated convolutional feature vectors for each image in our data set using the abovementioned CNN architectures. The convolutional feature maps can be considered as the compressed form of the input image data set, which can help the other learning algorithms to better understand the data set.

Fig. 2 Examples of feedstock images: **a** typical, **b** coarse particles, **c** feeding upset, and **d** obscured by dust. The center one-third of each image was retained for analysis
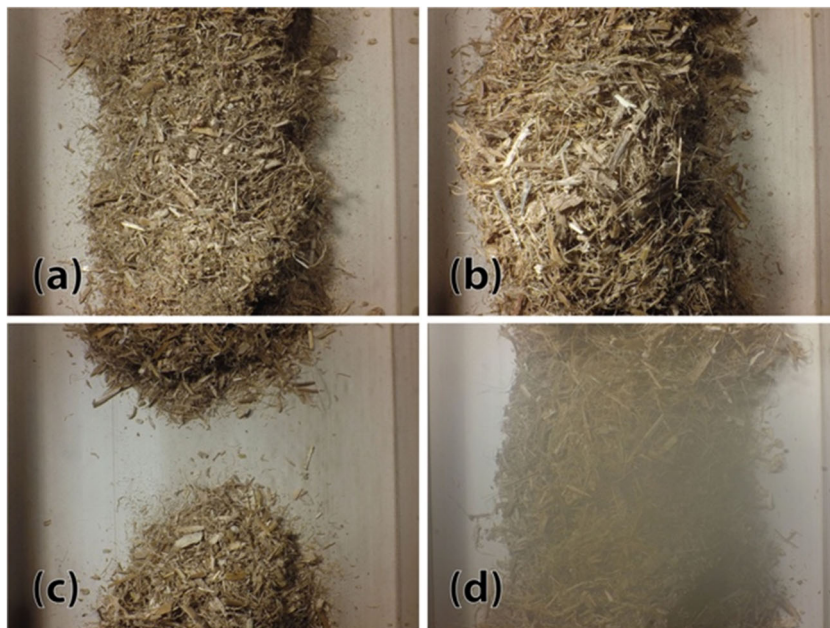
Table 1 Labeled image set

True images	7519
False images	18,604
Avg. true/false cycle period	28.0 min
Avg. anomaly (true) length	7.0 min

These convolutional feature vectors are then fed to fully connected NN layers to perform classification or regression operations, depending on the application. Several real-time applications, including Facebook's face detection and pedestrian detection [15], have demonstrated the usefulness of CNN layers and convolutional feature vectors rather than feeding raw pixel intensities to dense layers or fully connected layers to classify images [16]. Many image-clustering and image-classification challenges have been solved using feature vectors generated by neural networks that are pretrained on standard data sets like ImageNet or CIFAR [14, 16]. We used a similar approach in our application to identify anomalous feedstocks that may be leading to high motor loads in the cross-feeder. Figure 3 illustrates the basic process flow implemented for our application.

Stage 1. Convolutional neural network (CNN) feature vector generation. When we inspected the data set, we found that most of the images have feedstock distributed vertically toward the center third of the images. The leftmost and rightmost thirds of the images are dominated by the underlying weigh belt and shadows of the feedstock pile. The images were cropped to the central one-third as input for the CNN. We extracted feature vectors by passing these cropped input images to the CNN. We tried using several CNN architectures but obtained the best performance using ResNet (pretrained on the ImageNet data set) rather than using other, more standard, CNNs like VGG, LeNet, mobilenet, or Xception.

Stage 2. Principal component analysis (PCA). We faced a challenge of high dimensionality while handling feature vectors obtained from CNN. PCA was used to remove redundant features in the input data set and thereby enhance the training process [17]. Dimensionality was reduced while retaining 70%

of the variance of the input feature vectors. Table 2 illustrates the reduction in dimensionality using PCA for the tested CNN architectures.

Stage 3. Classification model. We needed to develop a classification model that could efficiently classify the PCA output obtained from stage 2. We initially tried classifying these feature vectors using several binary classification techniques like random forest classifier, logistic regression, and support vector machine (SVM). We also developed a time-series classifier using gated recurrent unit (GRU) neural networks [18]. These are recurrent neural networks whose architectures are very similar to long short-term memory (LSTM) networks. GRUs and LSTMs are advanced recurrent units that can perform better than traditional *tanh* units. Although GRUs and LSTMs generally give the same performance in most of the real-time applications, we chose a GRU network because the behavior of the gradient descent performance curves during the training process was more consistent (without any sudden fluctuations).

2.3 Pixel matrix feature parameterization

Individual images that contained objects other than biomass (i.e., air hose or operator hands) or were extensively obscured by dust or shadows were identified by visual inspection. Because these images did not lend themselves to image segmentation and analysis and would skew the results, they were removed from the data set for the subsequent analysis. In total, 878 images were removed based on multiple criteria (Table S1).

Individual images from all four pretreatment runs were analyzed using FIJI (ImageJ) [19]. The percent of the image covered in biomass was calculated for each image before it was cropped using thresholding that segmented the biomass from the background of the conveyor belt (Fig. 2). Textural features of cropped images were quantified using the plugins SurfCharJ 1q [20] and GLCM Texture Too [21]. The SurfCharJ 1q package calculated the root mean square deviation (Rq), arithmetical mean deviation (Ra), skewness (Rsk), kurtosis (Rku), lowest valley (Rv), highest peak (Rp), total profile height (Rt), mean height of surface profile (Rc), mean
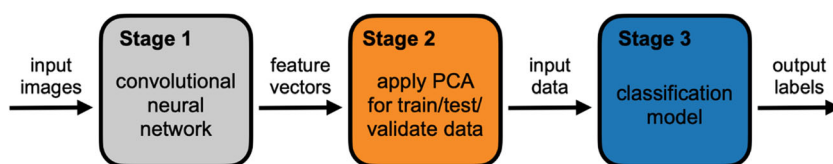
Fig. 3 Classification model workflow (PCA, principal component analysis)

Table 2 Output dimensions of several CNNs

CNN architecture	Shape of output feature vector	Feature vector size after PCA
VGG-16	(7,7,512)	(1,2555)
ResNet	(7,7,2048)	(1,3168)
Inception v3	(5,5,1536)	(1,488)
DenseNet	(7,7,1024)	(1,2238)
Xception	(7,7,2048)	(1,839)

polar facet orientation (FPO), variation of the polar facet orientation (MFOV), direction of azimuthal facets (FAD), mean resultant vector (MRV), and surface area (SA). The GLCM Texture Too calculated angular second moment, contrast, correlation, inverse difference moment, entropy, inertia, homogeneity, prominence, variance, and shade.

Some images still contained varying amounts of belt in the background even when cropped to a third of their original width, depending on the biomass distribution pattern. These distribution patterns were classified for each image into three types (based on the cropped images). The particle distribution patterns observed were (1) when the entire middle of the belt was covered in biomass because of its even distribution, referred to as "covered"; (2) when part of the middle third of the belt showed on either of both sides and/or corners, which occurred when the biomass bunched together, referred to as "clumped"; and (3) when gaps between the biomass occurred so that the belt was visible across the entire belt, including images that contained no biomass, referred to as "gaps".

Labeling each image according to these distribution patterns was of interest for two reasons. First, it enabled us to determine whether these distribution patterns were significant predictors of the TRUE/FALSE image labels. Second, these labels were included as dummy variables when analyzing the relationship of the textural features to the TRUE/FALSE variables to control for how these distribution patterns would affect the calculations of the textural features.

The textural data was analyzed using R statistical programming to determine if the textural features, percent area coverage, and/or distribution patterns were predictive of the TRUE/FALSE labels. Due to high multicollinearity found within the textural data, only angular second moment (ASM), prominence, shade, skewness (Rsk), mean height of the surface profile (Rc), MFOV, FAD, and MRV were analyzed, as they were sufficient to capture the variability of the rest of the data. Descriptive statistics were calculated and are included in the Supplemental Information (Table S2). The chi-squared test for independence was used to determine if a significant relationship between the biomass distribution types and TRUE/FALSE images existed. It was also used to determine if a significant relationship between the run number and the

TRUE/FALSE labels existed. Logistic regression analysis was used to determine if the textural features, percent area coverage, distribution patterns, and time were predictive of the TRUE/FALSE labels. The textural features were standardized to account for the range of their different magnitudes. Forward stepwise regression based on Akaike information criterion (AIC) values was used for variable selection to generate the best model when including the textural features. A categorical variable of the run number was included in the analysis to control for differences between data sets. A chi-squared test confirmed that the model including this categorical variable was statistically better than a model without it. A Wald chi-squared test was used to test for the significance of the variables in the logistic model generated by the stepwise function. The odds ratios and confidence intervals were calculated. The percent increase change in the odds for a one unit increase of each independent variable was also calculated based on standardized values. The level of statistical significance used was $p < 0.05$.

3 Results and discussion

3.1 Neural network models

After applying PCA on convolutional feature vectors obtained from the pretrained CNN, the next task was to develop a classification model that can pick anomalous feedstock from the given data set. We initially developed several binary classification models using random forests, logistic regression, and SVM. The confusion matrix of the model with random forest classifier in stage 3 of the workflow is shown in Fig. 3. Random forest classifier is an ensemble learning technique that takes advantage of a large number of individual decision trees [22]. Each individual tree in the random forest spits out a class prediction, and the class with the most votes becomes our model's prediction.

Because of the range and length of high motor torque time periods, up to 6 min instead of a short consistent interval of maybe 45 s, a simple binary classification method did not give reliable classification performance (Fig. 4a). We therefore migrated to using time-series classifiers using recurrent neural networks, namely, LSTM and GRU (Fig. 4b).

Quantile value used for automated image TRUE/FALSE labeling was also assessed to explore the sensitivity of the model performance. We initially used 70% quantile as a threshold value to pick instances of anomalous feedstock. We examined the model performance by labeling the data set using different quantile values (e.g., 60%, 65%, 75%, 80%). When the model is trained using 60% or 70% quantile-labeled images and tested using the same set, performance was the same in terms of model accuracy, precision, and recall (Fig. 5). Interestingly, if a model is trained using a

Fig. 4 Model performance confusion matrices with random forest binary classifier (**a**) and with GRU network time series classifier (**b**)
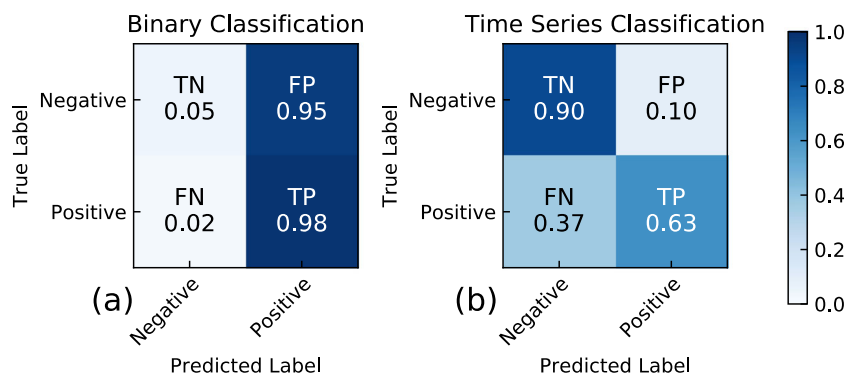
data set labeled with 60% quantile value and then tested on a data set labeled with 70% or 80% quantile, we observe a marked increase in model performance (Fig. 5).

A large number (approximately 1/8) of the total images were obscured from dust accumulation on the camera lens, which occurred during the latter half of the first of four runs before proper dust mitigation. In order to check the model sensitivity to the dusty images in our data set, we manually identified and removed dusty images and retrained the model from the edited subset. We observed no significant change in the model performance when the dusty images were removed, which suggests that the model is relatively insensitive to dust. This is likely because PCA applied in stage 2 discards constant features like the presence of dust in the input data set.

In real-time applications, it is important to alarm the operator or control system when anomalous feedstock has been detected. It should be acceptable to raise an alarm with a certain delay, depending on the application. Eventually, the important element is that the model should go to alarm state in the presence of anomalous feedstock with enough advance warning to avoid jamming process equipment such as the pug mill, cross-feeder, or plug-screw feeder, as in our application. This suggests that a negotiable delay in raising an alarm is likely acceptable, depending on the application. With this idea in mind, we evaluated model performance by considering false negatives as true positives for a certain

interval of time before and after the predicted alarm state. To explore this further, we employed thresholds as indicated in Fig. 6. A "left threshold" applies a negotiable amount of time before the alarm state in which false negatives can be considered as true positives. A "right threshold" applies a negotiable amount of time after the end of a predicted alarm state in which false negatives can be considered as true positives. Figure 6 illustrates how applying this technique greatly reduced false-negative results. The cyclical nature demonstrated in the motor loads plotted in Fig. 1b is a probable reason for this enhanced performance when a threshold is applied. The process we observed inherently occurs in a cycle with variable period and pulse width instead of a purely stochastic process. Therefore, this applied filter allows for higher discretion in model prediction and greater performance if the user of the results can handle a small delay on the order of a couple minutes.

3.2 Pixel matrix feature parameterization

Plots of textural features and percent area coverage of each image versus the time that the image was taken—with TRUE-labeled images shown in red and FALSE-labeled images show in gray—display the trends in image analysis-based features over the course of the biomass processing run (Fig. 7).
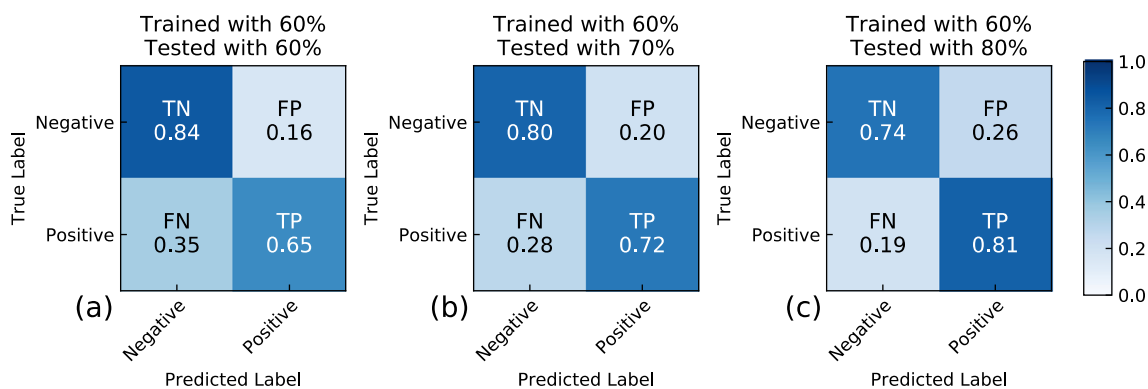
Fig. 5 Model performance sensitivity toward image true/false labeling quantile value. All models were trained with image labels determined above/below a cross-feeder torque quantile value of 60% and then tested with (**a**) images labeled using the same quantile value, (**b**) images labeled using a 70% quantile value, and (**c**) images labeled using an 80% quantile value
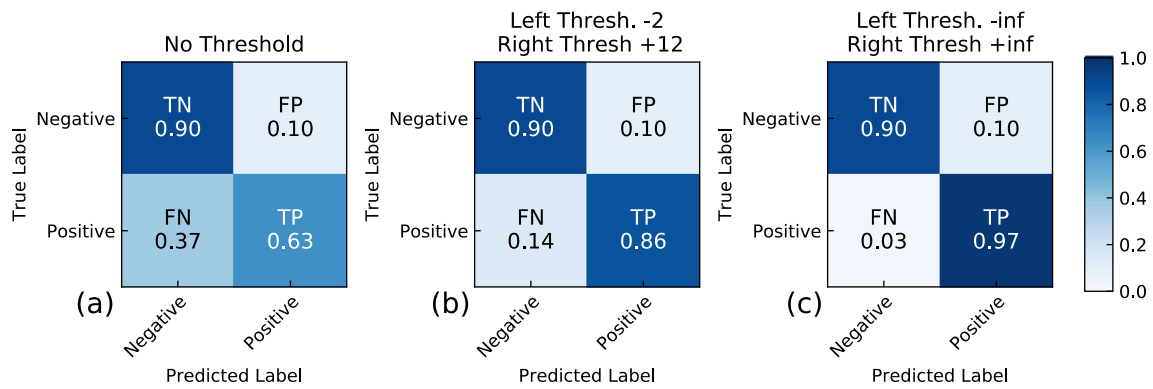
Fig. 6 Model performance when the occurrence of anomalous feedstock is generalized. (**a**) Native model prediction, (**b**) false-negatives considered true-positives before/after limited time step threshold range, (**c**) false-negatives considered true-positives across infinite time steps. Each time step is approximately 30 s

Chi-squared tests were first used to analyze if the distribution patterns and run numbers held significant relationships with the TRUE/FALSE labels to determine if these variables needed to be included in the logistic regression analysis. The distribution patterns of either the middle third of the belt being completely covered,
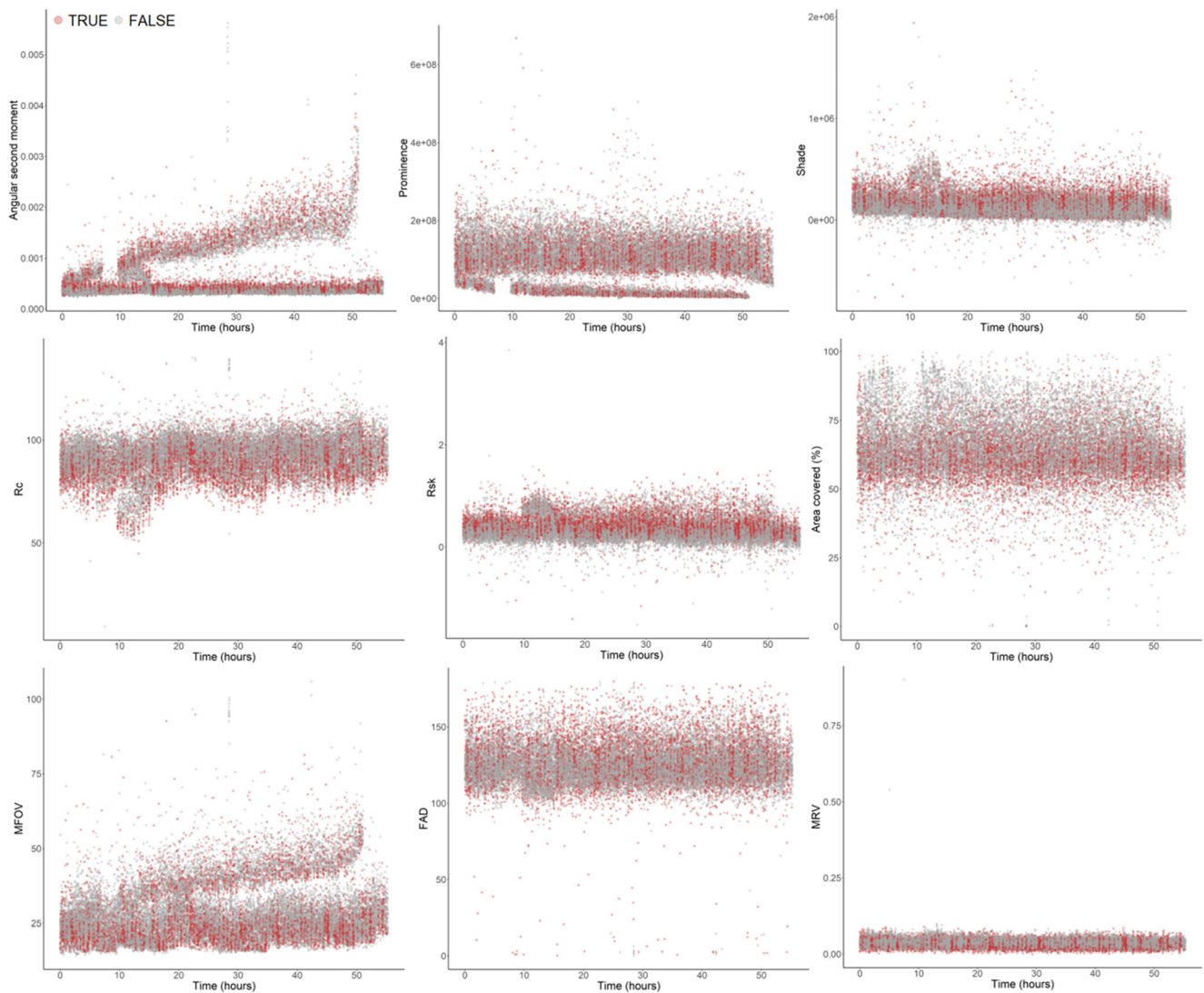

Fig. 7 Plots of textural features and percent area coverage of each image versus the time that the image was taken, with TRUE-labeled images shown in red and FALSE-labeled images shown in gray

showing clumps of biomass, or having gaps showing were found to have a significant relationship with the TRUE/FALSE label ($\chi^2 = 228$, df = 2, $p = 2.2e{-}16$) (Table S3). A marginally significant relationship existed between the TRUE/FALSE label and the run number ($\chi^2 = 6.73$, df = 3, $p = 0.081$) (Table S3). The run number was still included in the logistic regression analysis based on the observation that plots of the textural features versus time varied between runs (Fig. S1–S9).

Forward stepwise regression was used on the textural features of interest (angular second moment, shade, prominence, Rsk, Rc, MFOV, MRV, and FAD), time, the percent area covered by biomass, the distribution pattern categories, and run numbers. All these variables were included as significant in the model generated by the stepwise regression. The model coefficients for all variables except for time and run #2 were found to be significant (Table 3). Inclusion of categorical variables requires that one be used as a reference. The completely covered distribution pattern was used as a reference as well as run #4. Experimental run #4 was chosen as the reference after examining scatterplots of the textural features versus time for each run (Fig. S1–S9). A Wald chi-squared test determined that all explanatory variables except time were significant (Table S4). The model coefficients indicate that ASM, prominence, Rc, MRV, and percent area covered in biomass are negative predictors for the TRUE labels, while shade, Rsk, MFOV, and FAD are positive predictors. This means that increases in ASM, prominence, Rc, MRV, and percent area

are associated with a decrease in the odds of an issue in conveyance, while increases in shade, Rsk, MFOV, and FAD are associated with an increase in the odds of an issue in conveyance arising. It is worth noting that these relationships only hold true when all other variables are held constant.

The percent change in the odds of a TRUE label occurring per a change of one standardized unit was calculated based on the odds ratios that the regression model coefficients generated (Table 4). These results show that relative to when the belt is completely covered in biomass, the clumping distribution is about 26% less likely to be associated with causing an issue in conveyance, and the distribution pattern consisting of gaps is about 82% less likely. When examining the odds ratios based on the standardized textural features, the odds ratios indicate that increases of one standardized unit of angular second moment, prominence, Rc, and MRV are associated with about a 23%, 18%, 32%, and 26% decrease in odds of causing a TRUE label, respectively. The odds ratios indicate that shade, Rsk, MFOV, and FAD are associated with a 40%, 38%, 93%, and 10% increase in odds of causing a TRUE label, respectively. Every standardized unit increase of percent area coverage corresponds to a 41% decrease in odds of an issue in conveyance being caused. Example images of biomass on the weigh belt when these textural features were at the median of these textural parameters for their respective TRUE or FALSE labeled groups are shown in Fig. 8. While the images look similar overall, the TRUE examples display more stringy, high aspect ratio particles and more clumping, as evidenced by the shadowed pockets.

This study presents a novel finding that quantification and statistical analysis of textural features of images of piled biomass can offer insight into biorefinery operations such as the conveyance of biomass between units. The textural features measured from the images were found to be significant predictors of the TRUE and FALSE image labels, with angular second moment, prominence, Rc, and MRV being negative predictors and shade, Rsk, MFOV, and FAD all being positive predictors of anomalous biomass.

ASM is one of the better understood and broadly used textural characterization features. It is a measure of textural uniformity of an image that measures the local uniformity of gray levels; the more similar nearby pixels are, the larger ASM is. This indicates that biomass with high local uniformity is less likely to cause issues in conveyance. Shade and prominence each characterize the tendency of clustering of pixels in a region of interest [23] and are both measures of asymmetry, with higher values indicating more asymmetry. They are also both less used and less understood GLCM parameters [21, 24]. The asymmetric clustering that shade describes was found to be positively related to issues in conveyance, while that of prominence was found to be negatively related. This inverse relationship has been observed in previous studies [23].

Rc is the mean height of the surface profile elements, which in a grayscale image ranges from 0 to 255, with 0 being black

Table 3 Summary of model generated using standardized independent variables to account for their different orders of magnitude

	Est.	Standard error	z value	Pr9 > |z|	Significance
Intercept	−0.68	0.048	−14.3	$< 2 \times 10^{-16}$	***
Hour	0.00	0.00	1.59	0.112	*
Clumped	−0.29	0.05	−5.55	2.9×10^{8}	***
Gaps	−1.71	0.10	−17.1	$< 2 \times 10^{-16}$	***
Run #3	−0.47	0.05	−10.3	$< 2 \times 10^{-16}$	***
Run #2	−0.03	0.05	−0.53	0.594	*
Run #1	−0.47	0.10	−4.81	1.5×10^{-6}	***
ASM	−0.27	0.04	−6.01	1.8×10^{-9}	***
Prominence	−0.20	0.04	−4.47	7.7×10^{-6}	***
Shade	0.34	0.04	7.92	2.4×10^{-15}	***
Rsk	0.32	0.03	10.0	$< 2 \times 10^{-16}$	***
Rc	−0.39	0.05	−8.62	$< 2 \times 10^{-16}$	***
MFOV	0.65	0.07	10.0	$< 2 \times 10^{-16}$	***
FAD	0.09	0.02	5.62	1.9×10^{-8}	***
MRV	−0.30	0.02	−14.4	$< 2 \times 10^{-16}$	***
Area	−0.53	0.03	−20.4	$< 2 \times 10^{-16}$	***

$***p = 0$

$*p = 1$

All variables except time and run #2 have significant coefficients

Table 4 Percent change in odds of a TRUE label occurring per change in independent variables

	Percent change in odds of TRUE label per one standardized unit increase		
	Odds ratio	0.025%	0.975%
Intercept	− 49.6	− 54.1	− 44.6
Hour	0.2	0.0	0.4
Clumped	− 25.5	− 32.9	− 17.4
Gaps	− 81.9	− 85.1	− 78.0
Run #3	− 37.7	− 43.1	− 31.9
Run #2	− 2.5	− 11.1	7.0
Run #1	− 37.8	− 48.8	− 24.5
ASM	− 23.3	− 29.7	− 16.4
Prominence	− 18.1	− 25.1	− 10.7
Shade	40.4	29.1	52.7
Rsk	38.0	29.6	47.1
Rc	− 32.3	− 38.1	− 26.1
MFOV	92.5	69.4	118.8
FAD	9.5	6.1	13.1
MRV	− 25.6	− 28.5	− 22.5
Area	− 41.2	− 44.1	− 38.1

Distribution patterns of clumps and gaps are less likely to cause TRUE labels than if the center of the belt is completely covered. Increases in ASM, prominence, Rc, MRV, and percent area coverage will cause a decreased chance of a TRUE label, while increases in shade, Rsk, MFOV, and FAD are associated with an increased chance of a TRUE label

and 255 being white. The negative relationship that Rc has with the TRUE labels indicates that darker images are more likely to correlate with issues of poor conveyance. This same trend is indicated by the positive relationship that Rsk has with the TRUE labels, as positively skewed image matrixes result from images that have more dark pixels than light pixels. This association of darker images with issues in conveyance indicates that images with more extensive shadows in them are indicative of biomass loading patterns of concern. This makes sense because the images with large shadows in them occurred when larger clumps of biomass were present.

MFOV is a measurement of the variations in polar facet orientation of an image, which is a measure of the amount of variation in the z direction of the image as is indicated by the variations in gray levels [25]. The relative heights within an image are interpreted as a function of grayscale variation. The positive relationship between MFOV and the TRUE labels indicates that images with greater variation through the image of relative heights of pixels are related to issues in conveyance. This indicates that biomass distributions that are less uniform in the z plane but more disordered are more likely to cause issues in conveyance. FAD is a measurement of the direction of azimuthal facets, which represents the direction of local facets in the plane of the image [25]. MRV is a measure of the central tendency of the distribution of orientation direction, with values closer to zero indicating random distribution of orientations and values closer to 1 indicating that orientations of vectors are similar [26]. The negative relationship of

MRV with issues in conveyance indicates that when biomass is distributed randomly and without a more dominant gradient, it is less likely to cause issues with conveyance.

The discovery that percent area coverage of the belt has a negative relationship with TRUE labels was unexpected. This indicates that when the belt is completely covered, there is less of a chance of an issue in conveyance arising. However, the mass being conveyed on the belt is kept constant, perhaps indicating that either higher-density biomass distributions or biomass distributions that stick together such that they pile up instead of sideways could be responsible for issues in conveyance. This finding was contradicted by the fact that the one-third cropped images that showed the belt (either because of a gap or because a side/corner showed) were less likely to cause issues in conveyance than images that were completely covered with biomass. This possibly may have occurred because the percent surface area was calculated based on images that were not cropped; the reduction in conveyance problems were likely due to a short-lived reduction in feed rate and associated reduced motor loads.

An unexpected challenge that arose when analyzing the relationship between the full set of 23 textural features and the TRUE/FALSE images was a high degree of multicollinearity. Variables were eliminated with the goal of retaining as many as possible, and this presents its own challenge as several of the textural features left were ones that are less broadly used and less documented in the literature.

The lack of significance of the variable of time in predicting issues in conveyance stands out as particularly informative

Fig. 8 Examples of feedstock images from the central one-third of the field of view. These images show the morphology of the biomass pile in the conveyer belt when it classified as TRUE or FALSE. Images with median values for the textural parameters shade (**a**), Rsk (**b**), MFOV (**c**), and FAD (**d**) are shown, all of which had a positive correlation with increased likelihood of a strain on the cross-feeder

because this shows that the issues in conveyance have to do with the nature of the biomass itself (the textural features, how much of the belt it covers, and how evenly it is distributed on the belt). This highlights the utility for lignocellulosic biorefineries to understand the nature of—and measure in real time—the variability of the biomass material they handle.

4 Conclusions

Two methods of quantifying the handling quality of milled dry biomass feedstock using photographic images were developed and analyzed for performance. A neural network composed of convolutional and recurrent networks coupled to a fully connected binary classifier yielded the most favorable performance detecting anomalous feedstock, with accuracy reaching 63–97% true positive prediction depending on secondary application of a hysteresis filter. Pixel matrix feature parameterization using statistical methods applied to the 2D image pixel value vectors

offers complementary insight into what features are important in feedstock image classification. The textural features that were significant predictors of image labels were angular second moment, prominence, mean height of surface profile (Rc), mean resultant vector (MRV), shade, skewness (Rsk), variation of the polar facet orientation (MFOV), and direction of azimuthal facets (FAD). Expansion of these models is recommended across a wider variety of labeled feedstock images of different qualities and species to develop a more robust tool that may be deployed using low-cost cameras within biorefineries.

Code availability https://github.com/NREL/feedstock_machine_vision

Authors' contributions The manuscript was written through contributions of all authors. Chandrakanth Gudavalli and Elizabeth Bose are equal

first-author contributors. All authors have given approval to the final version of the manuscript.

Compliance with ethical standards

Conflicts of interest The authors declare that they have no conflict of interest.

Disclaimer The views expressed in the article do not necessarily represent the views of the DOE or the US Government. The US Government retains, and the publisher, by accepting the article for publication, acknowledges that the US Government retains a nonexclusive, paid-up, irrevocable, worldwide license to publish or reproduce the published form of this work, or allow others to do so, for US Government purposes.

Abbreviations ASM, angular second moment; CNN, convolutional neural network; FAD, direction of azimuthal facets; GRU, gated recurrent unit; LSTM, long short-term memory; MFOV, variation of the polar facet orientation; MRV, mean resultant vector; MV, machine vision; NN, neural network; NREL, National Renewable Energy Laboratory; PCA, principal component analysis; PMFP, pixel matrix feature parameterization; SA, surface area; SVM, support vector machine

References

1. U.S. Department of Energy Office of Energy Efficiency & Renewable Energy (2016). Biorefinery Optimization Workshop Summary Report. DOE/EE-1514; 2016. https://www.energy.gov/sites/prod/files/2017/02/f34/biorefinery_optimization_workshop_summary_report.pdf
2. Crawford NC, Nagle N, Sievers DA, Stickel JJ (2016) The effects of physical and chemical preprocessing on the flowability of corn stover. Biomass Bioenergy 85:126–134
3. Sievers DA, Kuhn EM, Thompson VS, Yancey NA, Hoover AN, Resch MG, Wolfrum EJ (2020) Throughput, reliability, and yields of a pilot-scale conversion process for production of fermentable sugars from lignocellulosic biomass: a study on feedstock ash and moisture. ACS Sustain Chem Eng 8(4):2008–2015
4. Yin S, Ding SX, Xie X, Luo H (2014) A review on basic data-driven approaches for industrial process monitoring. IEEE Trans Ind Electron 61(11):6414–6428
5. Boiarkina I, Depree N, Yu W, Wilson DI, Young BR (2017) Rapid particle size measurements used as a proxy to control instant whole milk powder dispersibility. Dairy Sci Technol 96(6):777–786
6. Harvill TL, Hoog JH, Holve DJ (1995) In-process particle size distribution measurements and control. Part Part Syst Charact 12(6):309–313
7. Wang S, Summers RM (2012) Machine learning and radiology. Med Image Anal 16(5):933–951
8. Lo SCB, Chan HP, Lin JS, Li H, Freedman MT, Mun SK (1995) Artificial convolution neural network for medical image pattern recognition. Neural Netw 8(7–8):1201–1214
9. Simonyan K, Vedaldi A, Zisserman A (2013) Deep inside convolutional networks: visualising image classification models and saliency maps. arXiv. https://arxiv.org/abs/1312.6034v1
10. Liu B, Yu X, Yu A, Wan G (2018) Deep convolutional recurrent neural network with transfer learning for hyperspectral image classification. J Appl Remote Sens 12(2):026028
11. Jmour N, Zayen S, Abdelkrim A (2018) Convolutional neural networks for image classification. In: international conference on advanced systems and electric technologies (IC_ASET), Hammamet, Tunisia, https://doi.org/10.1109/ASET.2018.8379889
12. Rawat W, Wang Z (2017) Deep convolutional neural networks for image classification: a comprehensive review. Neural Comput 29(9):2352–2449
13. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet classification with deep convolutional neural networks. NeurIPS. https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf
14. Liu S, Deng W (2015) Very deep convolutional neural network based image classification using small training sample size. In: 3rd IAPR Asian conference on pattern recognition (ACPR), Kuala Lupur, Malaysia, https://doi.org/10.1109/ACPR.2015.7486599
15. Szarvas M, Sakai U, Jun O (2006) Real-time Pedestrian Detection Using LIDAR and Convolutional Neural Networks. In: 2006 IEEE intelligent vehicles symposium, Tokyo, Japan, https://doi.org/10.1109/IVS.2006.1689630
16. Albawi S, Mohammed TA, Al-Zawi S (2017) Understanding of a convolutional neural network. In: international conference on engineering and technology (ICET), Antalya, Turkey, https://doi.org/10.1109/ICEngTechnol.2017.8308186
17. Wold S (1987) Principal component analysis. Chemom Intell Lab Syst 2(1–3):37–52. https://doi.org/10.1016/0169-7439(87)80084-9
18. Chung J, Gulcehre C, Cho K, Bengio Y (2014) Empirical evaluation of gated recurrent neural networks on sequence modeling. arXiv. https://arxiv.org/abs/1412.3555
19. Schindelin J, Arganda-Carreras I, Frise E, Kaynig V, Longair M, Pietzsch T, Preibisch S, Rueden C, Saalfeld S, Schmid B, Tinevez JY, White DJ, Hartenstein V, Eliceiri K, Tomancak P, Cardona A (2012) FIJI: an open-source platform for biological-image analysis. Nat Methods 9(7):676–682
20. Chinga G, Johnsen PO, Dougherty R, Berli EL, Walter J (2007) Quantification of the 3D microstructure of SC surfaces. J Microsc 227(3):254–265
21. Haralick RM (1979) Statistical and structural approaches to texture. Proc IEEE 67(5):786–804
22. Bosch A, Zisserman A, Muñoz X (2007) Image classification using random forests and ferns. In: Proceedings of IEEE International Conference on Computer Vision, Rio de Janeiro, Brazil, https://doi.org/10.1109/ICCV.2007.4409066
23. Aborisade DO, Ojo JA, Amole AO, Durodola AO (2014) Comparative analysis of textural features derived from GLCM for ultrasound liver image classification. Int J Comput Trends Technol 11:239–244. https://doi.org/10.14445/22312803/IJCTT-V11P151
24. Albregtsen F (2008) Statistical texture measures computed from gray level coocurrence matrices. Semantic Scholar. https://pdfs.

semanticscholar.org/3253/8c358410ebce7c9ecf688 addddf13f45b75b.pdf?_ga=2.219223705.1499168131. 1590080598-1607545833.1589490130

25. Chinga G, Gregersen Ø, Dougherty B (2003) Paper surface characterisation by laser profilometry and image analysis. J Microsc Anal 84:5–7

26. Curray JR (1956) The analysis of two-dimensional orientation data. J Geol 64(2):117–131. https://doi.org/10.1086/626329

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.