Ellipse Detection and Localization with Applications to Knots in Sawn Lumber Images

Shenyi Pan*
University of British Columbia
shenyi.pan@stat.ubc.ca

Shuxian Fan*
University of Washington
fansx@uw.edu

Samuel W.K. Wong
University of Waterloo
samuel.wong@uwaterloo.ca

James V. Zidek
University of British Columbia
jim@stat.ubc.ca

Helge Rhodin
University of British Columbia
rhodin@cs.ubc.ca

Abstract

While general object detection has seen tremendous progress, localization of elliptical objects has received little attention in the literature. Our motivating application is the detection of knots in sawn lumber images, which is an important problem since the number and types of knots are visual characteristics that adversely affect the quality of sawn lumber. We demonstrate how models can be tailored to the elliptical shape and thereby improve on general purpose detectors; more generally, elliptical defects are common in industrial production, such as enclosed air bubbles when casting glass or plastic. In this paper, we adapt the Faster R-CNN with its Region Proposal Network (RPN) to model elliptical objects with a Gaussian function, and extend the existing Gaussian Proposal Network (GPN) architecture by adding the region-of-interest pooling and regression branches, as well as using the Wasserstein distance as the loss function to predict the precise locations of elliptical objects. Our proposed method has promising results on the lumber knot dataset: knots are detected with an average intersection over union of 73.05%, compared to 63.63% for general purpose detectors. Specific to the lumber application, we also propose an algorithm to correct any misalignment in the raw lumber images during scanning, and contribute the first open-source lumber knot dataset by labeling the elliptical knots in the preprocessed images.

1. Introduction

Knots are formed by branches or limbs during the growth of a tree and commonly appear as dark ellipses on the surfaces of sawn lumber. As they significantly affect both the aesthetic quality and mechanical properties of lumber, knots have a central role in determining the commercial value of lumber. Such determination is performed by inspecting and measuring the visual characteristics on the surface of the piece. The cost associated with manual defect detection is prohibitive in industrial fabrication, and thus automated systems are needed. However, lumber is a highly variable material: sizes and shapes of knots and other defects, as well as color and texture of sawn lumber, have much variability from tree to tree. The automatic classification of sawn lumber is therefore more challenging than domains where computer vision algorithms are already applied routinely, such as for air bubble detection in glass or plastic casts.

Simple inspection systems that monitor the wood surface with color cameras are already deployed in the lumber industry, exploiting the fact that knots are usually of darker color than the background. Most commonly, color thresholding is used to mark all knot pixels with a '1'. However, this requires appropriate tuning either by a domain expert or threshold selection methods [26]. As a result, the performance of these detection methods is highly sensitive to the choice of threshold and associated geometric features selected for a particular wood type and camera setup. Moreover, ellipse detection in imperfectly binarized images remains a difficult and error-prone problem [19, 36, 3, 24, 35, 1]. We provide a comparison of our proposed deep-learning-based ellipse detection method to one of the geometric detection methods in [14] and to other learning-based variants on the lumber knot images in the experiment section.

In the literature, knots are typically modeled as elliptical cones when a piece of lumber is treated as a 3-dimensional object [13]. Hence, we model knot faces on a 2-dimensional surface are as ellipses (conic sections). The proposed knot detection method takes color images of lumber captured by high-definition cameras as the input and returns a 5-dimensional parameter vector, (cx, cy, rx, ry, θ), represent-

*Authors contributed equally.

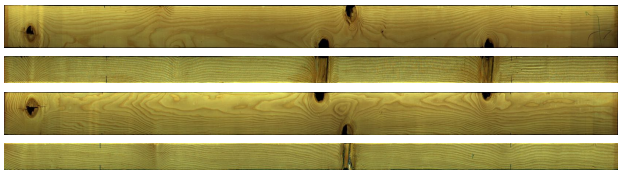

Figure 1: Sample lumber with the four surfaces scanned. The wide surfaces are shown on the first and the third rows and the narrow surfaces are shown on the second and last rows. Knots can be seen from their darker color and the noticeable grain distortion around them.

ing the position of the center of the knot face on the x-axis and y-axis, the length of the semi-diameters along the x- and y-axes, along with the rotation angle of the ellipse θ, respectively. The available data for this paper consist of images of 113 large sawn Douglas-fir lumber specimens with 4894 knots in total. The data are collected in a collaborative research project between the Department of Statistics at the University of British Columbia and FPInnovations. Figure 1 shows an example of knots on a piece of sawn lumber. Four images are taken for each board: two for the wide surfaces and another two for the narrow surfaces. For training and testing purposes, we manually labeled each knot.

In this paper, we address the knot detection and localization problem by adapting the Faster R-CNN framework [30] for rectangular object detection. Our main contributions are as follows:

- We adapt the Faster R-CNN to identify elliptical knots on the sawn lumber surfaces. Specifically, we replace the Region Proposal Network in the Faster R-CNN framework with the Gaussian Proposal Network to adapt to the ellipse detection problem, and propose an alternative loss function for ellipse regression.

- We propose an effective algorithm to fix the misalignment of raw images of lumber specimens.

- We prepare a labeled dataset of elliptical lumber knots, which can be used by future researchers to train and evaluate methods for knot detection and localization.

Knot detection is vital for assessing the strength and value of sawn lumber in a non-destructive manner. In particular, knot detection and localization based on lumber images is an integral step that generates input for existing knot matching algorithms that reconstruct the 3-dimensional structure of knots, e.g., [16]. Thus, this paper provides an important step in the pipeline towards advancing the state-of-the-art in automatic strength prediction of lumber and quality assessment of other materials containing elliptical defects.

This paper is organized as follows. Related work regarding lumber knot localization and elliptical object detection is reviewed in Section 2. The data preparation procedures and the lumber knot dataset are discussed in Section 3. Section 4 introduces our proposed method to solve the knot detection and localization problem. Experiment results and model performance evaluations are presented in Section 5. Section 6 concludes this paper.

2. Related Work

In this section, we discuss existing lumber defect detection methods, including active sensor-based solutions and vision-based solutions using geometry, machine learning (ML), and combinations thereof.

Active, using the laser tracheid effect. An existing method for knot location and size measurements in veneer (a thin layer of wood) surfaces using the laser tracheid effect is introduced in [34]. This method is based on the scattering patterns formed by laser spots. The laser light that penetrates the lumber surface is scattered mainly in the grain direction; this is often referred to as the "tracheid effect". The major features used to detect knots are the deviation of wood grain obtained by analyzing the amount of scattering. Areas with knots generally indicate the existence of large wood grain deviations and can be detected by thresholding. Methods based on the "tracheid effect" focus on finding improved approaches to computing the grain deviation. For example, Daval *et al.*[6] use the thermal conduction properties of lumber captured by a thermal camera to extract information such as the slope of grain and the presence of knots. The performance of the tracheid effect-based methods depends on the analysis of multiple individual laser spots projected on the surface. A practical limitation is in the resolution at which surface characteristics are captured. Empirical results show that tracheid effect-based methods are incapable of identifying knots of smaller sizes since they are usually located between adjacent laser spots. Therefore, these small knots are difficult to detect as they do not cause sufficiently large changes in the scattering patterns.

Passive geometric localization. Computer vision-based ellipse detection methods often use a multi-stage filtering process, by finding geometric features, such as lines, curves, arcs, and extended arc patterns as intermediate representations [19, 32, 17, 23, 4, 27, 5, 10, 28, 14, 7]. Improvements on the contour extraction and arc combination algorithms have been recognized in many real-world applications [36, 22, 15]. Previous work applied to knot detection used image processing, morphological processing, and feature extraction procedures to detect sizes and locations of knots on lumber surfaces [33, 38]. Global threshold-

ing is used to segment images through morphological operations to isolate regions that are likely to contain knots. Adaptive thresholding is then applied to suspected areas to improve the accuracy of knot segmentation results. However, the lack of robustness to noise and blurry shape edges still persists as major limitations to knot detection applications. The shape detection accuracy often deteriorates substantially when noise and blurriness of knots edges increase.

ML-based detection and localization. To identify and detect the locations of knots based on images, an alternative to modeling the knots as ellipses is to use object detection methods to find bounding boxes around knots. Bounding box regression and object detection have been a longstanding problem in the field of computer vision. The aim of bounding box regression is to refine or predict the minimum localization boxes within which objects of interest lie. With the developments in deep learning in recent years, many effective methods have been proposed to detect objects from an image. Methods such as R-CNN [12], Fast R-CNN [11], and Faster R-CNN [30] apply different methods to select possible regions containing objects. In particular, Faster R-CNN implements a Region Proposal Network (RPN) to propose plausible regions that are likely to contain objects. YOLO [29] segments images into smaller pieces and detects objects within and across the smaller pieces. SSD [21] is another single shot detection method similar to YOLO, but it uses feature maps from different layers to detect objects of different sizes. However, these object-detection methods are proposed to use bounding boxes to detect rectangular objects and have limitations when operating on elliptical objects with inherent symmetries.

Elliptical object localization. More recently, papers have been proposed to detect elliptical objects using machine learning-based methods. For example, Dong *et al.* [8] propose Ellipse R-CNN to detect clustered and occluded ellipses and suggest an objective function for performing ellipse regression. However, their method has a complicated pipeline which includes many image cropping and scaling operations tailored for detecting occluded ellipses, which is not the case for lumber knots. Moreover, Wang *et al.* [37] propose an Ellipse Proposal Network to detect optic boundaries in fundus images. However, the technical details regarding the objective functions and network architecture were not elaborated in the paper. Another application to pupil localization based on region proposal network and Mast R-CNN is proposed with a new computational schedule of anchor ellipse regression in [20]. Li [18] proposes a Gaussian Proposal Network (GPN) to replace the usual region proposal network in object detection frameworks such as Faster R-CNN. The elliptical proposals are parameterized as Gaussian distributions and the GPN is trained to minimize the Kullback-Leibler (KL) divergence between the ground truths and the proposed ellipses. However, the proposed GPN is only designed to be a replacement for the RPN and does not have the Region of Interest (RoI) pooling, classification, and regression components. Therefore, using GPN alone cannot complete the ellipse detection pipeline. Moreover, we improve the KL divergence loss by using the Wasserstein distance.

3. Data Preparation

In this section, we discuss the procedures for preparing the lumber knot dataset. Specifically, Section 3.1 introduces the proposed algorithm to fix the misalignment in the raw images, while Section 3.2 gives a brief introduction to the annotated lumber knot dataset.

3.1. Preprocessing Algorithm

Similar to the lumber image samples in Figure 1, we have a total of 113 pieces of lumber with four sides being scanned. The raw data do not contain any labels for knot faces, and the quality of the images is not ideal for object detection purposes. For example, it can be seen from Figure 1 that for each sample lumber image, there is one edge that is much longer than the other edge. Moreover, a major issue with the data is the pixel misalignment in the images of narrow surfaces.

An example of pixel misalignment in the image data is illustrated in Figure 2. To effectively improve the data quality, the preprocessing step needs to simultaneously recognize and address two challenges:

- The backgrounds do not have uniformly black color and contain noise pixels. Specifically, the dark background color exhibits an arbitrary change pattern close to the lumber edges.
- The color of knot faces can be very similar to the background. If a knot lies on either edge of a lumber board, it is often difficult to distinguish it from the background purely based on its color.

These features make it impractical to fix the misalignment issue by simply setting up a threshold to distinguish the background from the lumber area. Preliminary analysis shows that methods based on thresholds produce non-smooth edges of the lumber region and may fail when the knots lie on either edge of the surface.

Taking these considerations into account, we propose an iterative algorithm to fix the misalignment problem through column-by-column alignment of pixels. We illustrate the procedures in Figure 2. An RGB image is first converted to grey scale. For each pixel column, the algorithm determines its optimal shift by comparing the current column with its previous shifted neighbouring columns by taking

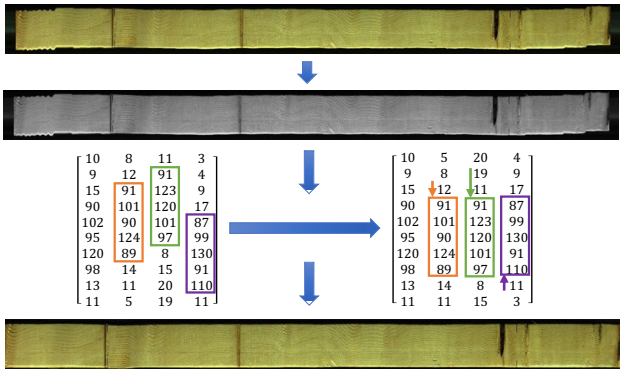

Figure 2: Illustration of fixing misalignment for one piece of lumber board. The RGB image is first converted to greyscale and fed to our algorithm to get the optimal displacement, which is used to shift the RGB image column-by-column to get the ideal result.

Figure 3: Bounding ellipse annotations for the visible knots on the wide surface of one sawn lumber.

the sum of the vector norms that are weighted inversely by the distances between columns. That is, the weights are proportional to the inverse of the distance raised to the power p. Denote the pixel values in the ith column by $\boldsymbol{c}_i$. Let the number of neighbours be n and the norm order be k. The optimal shift $\widehat{s}$ for column i, $i > 1$, is obtained by searching among all possible shifts $s \in \{s^i_{\min}, \ldots, -1, 0, 1, \ldots, s^i_{\max}\}$ such that

$$\widehat{s} = \underset{s \in \{s^i_{\min}, \ldots, -1, 0, 1, \ldots, s^i_{\max}\}}{\arg\min} \sum_{j=\min(1, i-n)}^{i-1} \frac{1}{|i-j|^p} ||\boldsymbol{c}^s_i - \boldsymbol{c}_j||_k, \tag{1}$$

where $\boldsymbol{c}^s_i$ represents pixels of the ith column shifted by s.

The values of n, p, and k are fine-tuned to achieve optimal results. Taking the extent of misalignment into account, we set $n = 100$, $k = 2$, and $p = 1$ to search for the optimal shifts for each pixel column. This method proves to be effective in resolving the misalignment in the lumber images. A simplified example is depicted in Figure 2. It can be seen that our proposed greedy method is robust to the nonuniform dark background and edge-lying knots.

3.2. Lumber Knot Dataset

After removing the dark background, lumber images are manually annotated using the free image annotation VGG Image Annotator [9] with each ellipse parameterized by five parameters: the x and y coordinates of the center of each ellipse denoted by cx and cy; the semi-diameters along the x- and y-axes denoted by rx and ry; the counterclockwise rotation angle in radians denoted by θ.

Figure 3 shows an example of the bounding ellipses for all the visible knots on the wide surface of a piece of lumber. It can be seen that all the visible knots on the board are accurately and tightly annotated. We then randomly crop and resize the annotated images to generate square images for the detection task. This step also re-computes the parameters of the bounding ellipse based on their relative positions in each cropped image. The dimension of each cropped and resized image is 512×512 pixels. For each piece of lumber, an average of 57.4 cropped images that partially contain at least one elliptical knot are generated. The complete dataset contains 4894 annotated lumber knots and is freely accessible on `https://forestat.stat.ubc.ca/tiki-index.php?page=Databases+-+Products` to the public for future research.

4. Proposed Method

In this section, we formulate our elliptical detection and localization method. The input to our method is a single image. The outputs are detected ellipses parameterized by (cx, cy, rx, ry, θ). Since we are interested in a single class of knots, there is no need to further classify the category to which the objects contained in the bounding ellipses belong. Our approach combines Faster R-CNN and its region proposal networks with the GPN introduced in [18]. The overview of the model pipeline is discussed in Section 4.1. We introduce the GPN in Section 4.2. Our extensions in terms of the region proposal branch and ellipse regression loss functions are explained in Sections 4.3 and 4.4, respectively.

4.1. Overview of the Model Pipeline

As an illustration, Figure 4 shows the architecture of our proposed model. We adopt the basic architecture of the Faster R-CNN, which was originally designed to solve rectangular object detection problems. An image is first passed to the convolutional layers for feature map extraction. Instead of proposing bounding boxes, we use the GPN to propose bounding ellipses as 2D equi-probability contours of Gaussian distributions on the image plane. With the RoI pooling branch, the feature maps corresponding to the proposed regions are then obtained. The feature maps and proposals are finally fed to the ellipse regression and classification branch for final ellipse prediction.

4.2. Gaussian Proposal Network

In this section, we introduce the adaptations made by the GPN to accommodate Faster R-CNN for localizing ellipses, as required for the lumber knot detection problem.

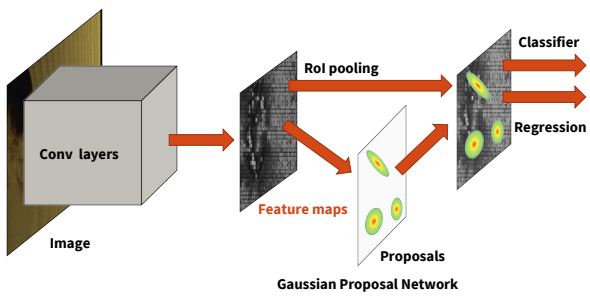

Figure 4: Overview of the ellipse localization and prediction model pipeline.

4.2.1 Parameterizing Ellipses by Gaussian Distributions

In [18], ellipses are reparameterized as 2-dimensional contours of Gaussian distributions. As a result, the usual objective function for minimizing the L1 or L2 loss used in performing bounding box regression can be replaced by minimizing the distance metrics between two Gaussian distributions for ellipse regression. This section shows how ellipses can be represented using 2D Gaussian distributions.

An ellipse in a 2D coordinate system without rotation can be represented by

$$\frac{(x-\mu_x)^2}{\sigma_x^2} + \frac{(y-\mu_y)^2}{\sigma_y^2} = 1, \tag{2}$$

where μ_x and μ_y are coordinates representing the centers of the ellipse, and σ_x and σ_y are the lengths of the semi-axes along the x and y axes.

The probability density function of a 2D Gaussian distribution is given by

$$f(\boldsymbol{x}|\boldsymbol{\mu},\boldsymbol{\Sigma}) = \frac{\exp\left(-\frac{1}{2}(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu})\right)}{2\pi|\boldsymbol{\Sigma}|^{\frac{1}{2}}}, \tag{3}$$

where the vector $\boldsymbol{\mu}$ denotes the coordinate vector representing the center (x, y), while $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ are the mean vector and covariance matrix of the Gaussian distribution. If we assume the off-diagonal entries of $\boldsymbol{\Sigma}$ are 0 and parameterize $\boldsymbol{\mu}$ and $\boldsymbol{\Sigma}$ as

$$\boldsymbol{\mu} = \begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix} \text{ and } \boldsymbol{\Sigma} = \begin{bmatrix}\sigma_x^2 & 0\\ 0 & \sigma_y^2\end{bmatrix}, \tag{4}$$

the equation in Equation (2) for the ellipse corresponds to the density contour of the 2D Gaussian distribution when

$$(\boldsymbol{x}-\boldsymbol{\mu})^T\boldsymbol{\Sigma}^{-1}(\boldsymbol{x}-\boldsymbol{\mu}) = 1. \tag{5}$$

When the major axis of the ellipse is rotated of an angle θ with respect to the x-axis, a rotation matrix $R(\theta)$ can be defined as

$$R(\theta) = \begin{bmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{bmatrix}. \tag{6}$$

This matrix can be used to map the coordinates in the original (x, y) system into a new (x', y') system, i.e.,

$$\begin{bmatrix}x'\\ y'\end{bmatrix} = R(\theta)\begin{bmatrix}x\\ y\end{bmatrix} \tag{7}$$

Denote the lengths of the semi-major and semi-minor axes of the ellipse by σ_l and σ_s. It can be shown that the ellipse centered at (μ_x, μ_y) with semi-major and semi-minor axes of lengths σ_l and σ_s, and a rotation angle of θ between its major axis and the x-axis where $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ can be parameterized by a 2D Gaussian distribution with

$$\boldsymbol{\mu} = \begin{bmatrix}\mu_x\\ \mu_y\end{bmatrix} \text{ and } \boldsymbol{\Sigma}^{-1} = R^T(\theta)\begin{bmatrix}1/\sigma_l^2 & 0\\ 0 & 1/\sigma_s^2\end{bmatrix}R(\theta). \tag{8}$$

4.2.2 Replacing Region Proposal Network with Gaussian Proposal Network

The goal of the GPN is to propose bounding ellipses such that the Gaussian parameters $(\mu_x, \mu_y, \sigma_l, \sigma_s, \theta)$ are close to the ground truth ellipses through a distance metric. In [18], Kullback-Leibler (KL) divergence is used as the distance measure. It is easily seen that the KL divergence between a proposed 2D Gaussian distribution $\mathcal{N}_p(\boldsymbol{\mu}_p, \boldsymbol{\Sigma}_p)$ and a target 2D Gaussian distribution $\mathcal{N}_t(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ is

$$D_{KL}(\mathcal{N}_p||\mathcal{N}_t) = \frac{1}{2}[\mathrm{tr}(\boldsymbol{\Sigma}_t^{-1}\boldsymbol{\Sigma}_p) + (\boldsymbol{\mu}_p-\boldsymbol{\mu}_t)^T\boldsymbol{\Sigma}_t^{-1}(\boldsymbol{\mu}_p-\boldsymbol{\mu}_t) + \log\frac{|\boldsymbol{\Sigma}_t|}{|\boldsymbol{\Sigma}_p|} - 2], \tag{9}$$

where $\mathrm{tr}(\cdot)$ is the trace of a matrix. GPN replaces RPN in the Faster R-CNN framework by predicting five parameters for ellipses instead of four parameters for bounding boxes and minimizing KL divergence instead of the smooth L1 loss. With these modifications, the RPN module can be replaced by the GPN to propose bounding ellipses.

4.3. Region Proposal and Offset Regression

GPN was originally only designed to replace the RPN in the Faster R-CNN framework to generate elliptical proposals with the remaining components removed. Further detecting and predicting the exact locations of the ellipses in an image is out of the scope of GPN. We re-introduce the necessary RoI pooling as well as the ellipse classification and regression components to complete the ellipse detection pipeline in analogy to Faster R-CNN. Faster R-CNN pools the feature maps in a rectangular region around the detection to have a fixed resolution, applies a CNN to these, and predicts offsets between the proposed parameters and the ground-truth parameters for each detection using a fully-connected layer. Similarly, given each ellipse proposal output by the GPN, we find the tightest axis-aligned bounding

box covering the proposal. Feature maps in this rectangular are pooled in a similar manner to the Faster R-CNN. A fully-connected layer is then used to predict the parameter offsets based on the proposal center as well as the ellipse direction and shape for each detected ellipse.

4.4. Loss Function for Ellipse Regression

In Faster R-CNN, L1 or L2 losses are used as the loss function to predict the offsets between the four pairs of predicted and ground-truth parameters defining a rectangular object. However, these losses do not work well for elliptical object detection problems since the angle parameter needs to be treated differently than the other four parameters.

A natural choice for the ellipse regression loss is the KL divergence in GPN, which is used as the distance measure between two Gaussian distributions in generating ellipse proposals. Nevertheless, KL divergence has a few non-negligible drawbacks. Firstly, KL divergence is asymmetric, i.e., $D_{KL}(\mathcal{D}_1||\mathcal{D}_2)$ does not always equal $D_{KL}(\mathcal{D}_2||\mathcal{D}_1)$ given two distributions $\mathcal{D}_1$ and $\mathcal{D}_2$. Secondly, KL divergence can sometimes be numerically unstable and it tends to be very large in magnitude when the two distributions are far apart. This may cause problems in gradient evaluation, hindering convergence of the neural network. Lastly, KL divergence does not satisfy the triangle inequality and is not a valid mathematical metric in that sense.

Wasserstein distance is another popular distance measure defined between two probability distributions. In contrast to KL divergence, Wasserstein distance is symmetric and satisfies the triangle inequality. In recent years, Wasserstein distance has been proposed to replace other asymmetric losses to improve the results generated by neural network models. For example, Arjovsky *et al.* [2] propose Wasserstein GAN to stabilize the training of GANs.

The p-Wasserstein distance between two probability measures μ and ν is

$$W_p(\mu,\nu) = \left(\inf \mathbb{E}[d(X,Y)^p]\right)^{1/p}, \tag{10}$$

where $d(\cdot,\cdot)$ is a norm function, $X \sim \mu$, and $Y \sim \nu$. For two 2D Gaussian distributions, the 2-Wasserstein distance between them with respect to the usual Euclidean norm has a convenient form. For a proposed 2D Gaussian distribution $\mathcal{N}_p(\boldsymbol{\mu}_p, \boldsymbol{\Sigma}_p)$ and a target 2D Gaussian distribution $\mathcal{N}_t(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$, the 2-Wasserstein distance with respect to the Euclidean norm is

$$[W_2(\mathcal{N}_p,\mathcal{N}_t)]^2 = ||\boldsymbol{\mu}_p - \boldsymbol{\mu}_t||_2^2 + \mathrm{tr}\Big[\boldsymbol{\Sigma}_p + \boldsymbol{\Sigma}_t - 2\left(\boldsymbol{\Sigma}_p^{\frac{1}{2}}\boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_p^{\frac{1}{2}}\right)^{\frac{1}{2}}\Big] \tag{11}$$

according to the results in [25], where $\mathrm{tr}(\cdot)$ is the trace of a matrix. In the commutative case where $\boldsymbol{\Sigma}_p\boldsymbol{\Sigma}_t = \boldsymbol{\Sigma}_t\boldsymbol{\Sigma}_p$, the 2-Wasserstein distance between two 2D Gaussian distributions can be further simplied to

$$[W_2(\mathcal{N}_p,\mathcal{N}_t)]^2 = ||\boldsymbol{\mu}_p - \boldsymbol{\mu}_t||_2^2 + ||\boldsymbol{\Sigma}_p^{\frac{1}{2}} - \boldsymbol{\Sigma}_t^{\frac{1}{2}}||_F^2, \tag{12}$$

where $||\cdot||_F$ is the Frobenius norm of a matrix. The two covariance matrices can be computed based on the inverses in Equation (8) from the ellipse parameters. The square root matrices $\boldsymbol{\Sigma}_p^{\frac{1}{2}}$ and $\boldsymbol{\Sigma}_t^{\frac{1}{2}}$ have a closed-form solution:

$$\boldsymbol{\Sigma}_p^{\frac{1}{2}} = R_p^T(\theta)\begin{bmatrix}\sigma_{p,l} & 0\\ 0 & \sigma_{p,s}\end{bmatrix}R_p(\theta) \tag{13}$$

and

$$\boldsymbol{\Sigma}_t^{\frac{1}{2}} = R_t^T(\theta)\begin{bmatrix}\sigma_{t,l} & 0\\ 0 & \sigma_{t,s}\end{bmatrix}R_t(\theta). \tag{14}$$

With all these modifications and adaptions, the overall loss function is the weighted sum of three components: the GPN ellipse proposal loss, the ellipse regression loss, and the cross entropy of classifying ellipses and backgrounds.

5. Experimental Results

Comprehensive evaluations of our proposed method for detecting elliptical knots in the lumber knot dataset are presented in this section. We introduce the experimental setup in Section 5.1 and summarize the quantitative detection performances across different experiment settings in Section 5.2. Visualizations of the detection results are provided in Section 5.3. We also compare our proposed method against the baseline deep learning-based solution and geometric reconstructions in 5.4.

5.1. Experiment Setup

Among all the annotated lumber specimens in the lumber knot dataset, 70% are randomly chosen as the training set, 10% are chosen as the validation set, and the remaining 20% are used as the test set. We trained the model on the training set for 20 epochs; the model with the lowest validation total loss is saved for testing purposes. As with the original Faster R-CNN model, the pretrained VGG-16 network in [31] is used as the base model for ellipse proposal generation and feature map extraction. We trained our proposed model with PyTorch 1.0.

The average of the intersection over union (IoU) between all the detected ellipses and ground truth ellipses are computed and used as the metrics to evaluate the performance of a proposed detection methods. Since there are not closed-formed formula to compute the overlapping area between two ellipses, we draw a grid of points from both ellipses and use a discretized sampling method to compute the IoU between the two ellipses. The width and height of the point grid are the same as those of the tightest axis-aligned rectangle covering both ellipses. One point is assigned at each pixel location.

Setting	RPN, L2	GPN, L2	GPN, KLD	GPN, 2-WD
Exp 1	60.16	64.74	73.37	73.00
Exp 2	65.59	66.14	73.13	73.32
Exp 3	64.54	65.82	72.48	72.98
Exp 4	65.98	66.97	71.07	72.85
Exp 5	61.87	65.40	73.26	73.10
Mean IoU	63.63	65.81	72.66	73.05
s.e.	2.52	0.83	0.95	0.18

Table 1: The IoUs between the detected and ground-truth ellipses under each of the four settings in each repeated experiment along with the mean IoU and the standard error. L2, KLD, and 2-WD represent L2 loss, KL divergence, and 2-Wasserstein distance, respectively. All numbers are in percentages.

5.2. Quantitative Evaluation of Detection Performance

To evaluate the performance of our proposed method, four experimental settings are considered as follows:

1. *RPN with L2 loss for ellipse offset regression.* RPN in the Faster R-CNN is used to generate bounding ellipse proposals. Instead of the four parameters characterizing a rectangle, the RPN outputs the five parameters characterizing an ellipse. L2 loss is used for the ellipse offset regression. This setting directly modifies the RPN for the ellipse detection problem and serves as the baseline model.

2. *GPN with L2 loss for ellipse offset regression.* RPN in the Faster R-CNN is replaced with the GPN to generate ellipse proposals. L2 loss is used for ellipse offset regression.

3. *GPN with KL divergence for ellipse offset regression.*

4. *GPN with 2-Wasserstein distance for ellipse offset regression.*

Note that GPN uses KL divergence while RPN uses smooth L1 loss in the proposal network. We trained the model five times under each setting. The IoUs between the detected and ground-truth ellipses under each setting in each repeated experiment along with the mean IoU and standard error are reported in Table 1.

From Table 1, it can be seen that using our proposed method, which generates elliptical proposals with GPN and uses the 2-Wasserstein distance as the loss function for ellipse offeset regression, improves the mean IoU by 10% compared with a general detector such as Faster R-CNN, which generates proposals with RPN and uses L2 distance as the loss function for offset regression. Directly replacing RPN in Faster R-CNN with GPN improves the mean IoU by around 2%. Replacing the general-purpose L2 loss with KL divergence or 2-Wasserstein distance both significantly improves the ellipse detection performance. In particular, models using 2-Wasserstein distance outperform models using KL divergence by more than 0.4%. Moreover, using the Wasserstein distance instead of KL divergence led to a much lower (0.18 vs. 0.95) standard deviation across experiments.

5.3. Qualitative Evaluation of Detection Performance

In the previous section, we quantitatively evaluated the performance of our proposed ellipse detection method against the baseline Faster R-CNN model. In this section, visualizations of the detected ellipses in the lumber knot dataset are provided to qualitatively assess our proposed method.

Figure 5 shows the ground-truth knots from three pieces of lumber as well as the detected knots using the baseline method (RPN, L2) and our proposed method (GPN, 2-WD). It can be seen that the elliptical bounds for knots are drawn more tightly and accurately using our proposed method compared with the baseline method. This holds particularly true for knots close to the board boundary, which are only partially visible. Nonetheless, it can be seen that the rotation angles of the detected ellipses have relatively large errors in some of the examples shown in Figure 5. Future research can be done to further improve the prediction performance for the rotation angles, possibly by introducing oriented instead of axis-aligned pooling regions in the GPN.

5.4. Comparisons to Geometric Ellipse Detection

To compare with our proposed method, we applied the geometric fast ellipse detector using projective invariant pruning method proposed in [14] to the lumber knot dataset (their code is available on `https://github.com/dlut-dimt/ellipse-detector`). Jia's method uses geometric features to find ellipses, which performs extremely poorly on lumber knot images. Among all the test images, less than 1% of the ellipses can be detected using their method. In particular, Jia's method failed to detect any ellipses in all the 12 images in Figure 5. Furthermore, Jia's method is also sensitive to the positioning of ellipses. For example, Figure 6 visualizes the detection results using our proposed method (GPN + 2-WD) versus Jia's method on the same knot across different cropped images of a lumber board. It can be seen that among these five positions, our method consistently generates accurate position estimations while Jia's method is only able to successfully yet inaccurately detect this ellipse at one out of the five positions. Therefore, we conclude that our method is more robust and

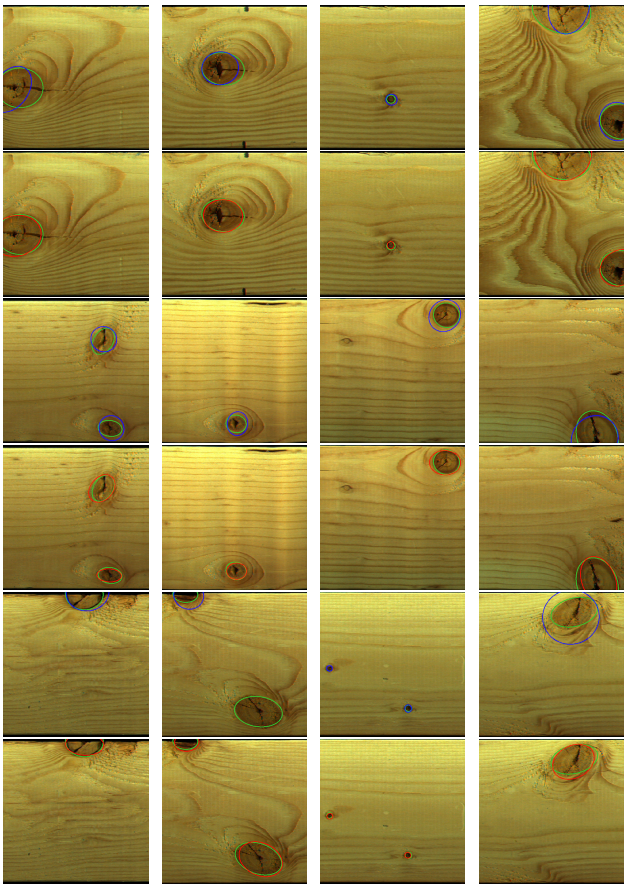

Figure 5: Examples of detected ellipses from three specimens of lumber using the baseline method (RPN, L2) and our proposed method (GPN, 2-WD). Images in every two rows are from the same specimen. Green, blue, and red ellipses are the ground-truth ellipses, the detected ellipses using the baseline method, and the detected ellipses using our proposed method, respectively.

reliable than non-machine learning-based ellipse methods and works much better on detecting knots in lumber images. In terms of runtime, for an image of 512×512 pixels, Jia's method takes around 8.8 milliseconds to generate predictions, while our proposed method (GPN + 2-WD) takes around 226 milliseconds in the test stage. Although Jia's method is faster than our proposed method, both methods are sufficiently fast for real-world applications.

6. Conclusion

In this paper, we propose a method tailored to detect and localize elliptical objects in images. Our method adapts the Region Proposal Network in Faster R-CNN to model elliptical objects. We also extend the existing Gaussian Proposal Network by adding RoI pooling as well as the ellipse classification and regression branches to complete the elliptical object detection pipeline. Furthermore, we propose Wasserstein distance as the loss function for ellipse offset predictions. Experiments on the sawn lumber images show that our proposed method improves the mean IoU of the detected ellipses by 10% compared with the baseline method. This is a major improvement in this illustrative application in so far as the result will be used in models that predict lumber strength and hence determine the grade class into which a piece of lumber will be placed. That in turn benefits consumers who use these grades in selecting lumber for their specific application. And it benefits producers who will use machine grading techniques to classify their lumber. For the forest products industry, this can result in products of better quality for intended use and in turn improvements to the manufacturer's bottom line. In addition, specific to the lumber example, we propose an algorithm to correct the misalignment of raw images of lumber specimens and create the first open-source lumber knot dataset with labeled elliptical knots, which can be used for future research. While experiments in this paper focus on the knot detection problem in sawn lumber images, our proposed method can easily be applied to detect ellipses in datasets containing other types of elliptical objects. In future work, we will attempt to predict the 3D knot shapes given 2D ellipsoid supervisions from both sides. It would also be interesting to predict the entire elliptical cone given images of all three sides, supervised by the 2D ellipse/Gaussian formulation.

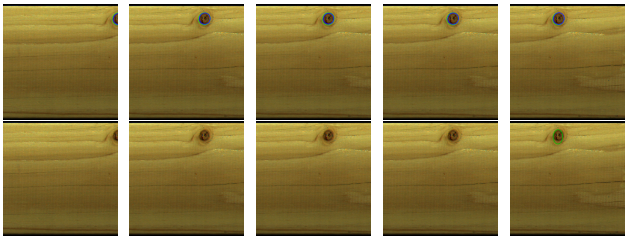

Figure 6: Examples of detecting the same ellipse in different cropped images using our method (GPN, 2-WD) and Jia's method. Green and blue ellipses in the first row are the ground-truth ellipses and the detected ellipses using our proposed method. Green ellipses in the second row are the detected ellipses using Jia's method. Note that Jia's method only detects the knot in the last cropped image.

Acknowledgements. The authors thank Conroy Lum, along with FPInnovations and its technical support staff, for facilitating the experimental work that was done to produce the data used in this paper. We also thank Terry Chen, Zongjun Liu, Angela Wang-Lin, and Caitlin Zhang for their assistance in data processing. The work reported in this paper was partially supported by FPInnovations and a Collaborative Research and Development grant from the Natural Sciences and Engineering Research Council of Canada.

References

[1] AS Aguado, ME Montiel, and MS Nixon. Ellipse detection via gradient direction in the hough transform. In *Fifth International Conference on Image Processing and its Applications, 1995.*, pages 375–378. IET, 1995.

[2] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *International Conference on Machine Learning*, pages 214–223, 2017.

[3] Alex Yong Sang Chia, Maylor KH Leung, How-Lung Eng, and Susanto Rahardja. Ellipse detection with hough transform in one dimensional parametric space. In *2007 IEEE International Conference on Image Processing*, volume 5, pages V–333. IEEE, 2007.

[4] Alex Yong-Sang Chia, Susanto Rahardja, Deepu Rajan, and Maylor Karhang Leung. A split and merge based ellipse detector with self-correcting capability. *IEEE Transactions on Image Processing*, 20(7):1991–2006, 2010.

[5] Alex Yong-Sang Chia, Deepu Rajan, Maylor Karhang Leung, and Susanto Rahardja. Object recognition by discriminative combinations of line segments, ellipses, and appearance features. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(9):1758–1772, 2011.

[6] Vincent Daval, Guillaume Pot, Mohamed Belkacemi, Fabrice Meriaudeau, and Robert Collet. Automatic measurement of wood fiber orientation and knot detection using an optical system based on heating conduction. *Optics Express*, 23(26):33529–33539, 2015.

[7] Huixu Dong, Dilip K Prasad, and I-Ming Chen. Accurate detection of ellipses with false detection control at video rates using a gradient analysis. *Pattern Recognition*, 81:112–130, 2018.

[8] Wenbo Dong, Pravakar Roy, Cheng Peng, and Volkan Isler. Ellipse r-cnn: Learning to infer elliptical object from clustering and occlusion. *arXiv preprint arXiv:2001.11584*, 2020.

[9] Abhishek Dutta and Andrew Zisserman. The VIA annotation software for images, audio and video. In *Proceedings of the 27th ACM International Conference on Multimedia*, MM '19, New York, NY, USA, 2019. ACM.

[10] Michele Fornaciari, Andrea Prati, and Rita Cucchiara. A fast and effective ellipse detector for embedded vision applications. *Pattern Recognition*, 47(11):3693–3708, 2014.

[11] Ross Girshick. Fast r-cnn. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1440–1448, 2015.

[12] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 580–587, 2014.

[13] Pablo Guindos and Manuel Guaita. A three-dimensional wood material model to simulate the behavior of wood with any type of knot at the macro-scale. *Wood Science and Technology*, 47(3):585–599, 2013.

[14] Qi Jia, Xin Fan, Zhongxuan Luo, Lianbo Song, and Tie Qiu. A fast ellipse detector using projective invariant pruning. *IEEE Transactions on Image Processing*, 26(8):3665–3679, 2017.

[15] Ren Jin, Hafiz Muhammad Owais, Defu Lin, Tao Song, and Yifang Yuan. Ellipse proposal and convolutional neural network discriminant for autonomous landing marker detection. *Journal of Field Robotics*, 36(1):6–16, 2019.

[16] Seong-Hwan Jun, Samuel WK Wong, James V Zidek, Alexandre Bouchard-Côté, et al. Sequential decision model for inference and prediction on nonuniform hypergraphs with application to knot matching from computational forestry. *The Annals of Applied Statistics*, 13(3):1678–1707, 2019.

[17] Euijin Kim, Miki Haseyama, and Hideo Kitajima. Fast and robust ellipse extraction from complicated images. In *Proceedings of IEEE Information Technology and Applications*. Citeseer, 2002.

[18] Yi Li. Detecting lesion bounding ellipses with gaussian proposal networks. In *International Workshop on Machine Learning in Medical Imaging*, pages 337–344. Springer, 2019.

[19] Lars Libuda, Ingo Grothues, and Karl-Friedrich Kraiss. Ellipse detection in digital image data using geometric features. In *Advances in Computer Graphics and Computer Vision*, pages 229–239. Springer, 2007.

[20] Horng-Horng Lin, Zheng-Yi Li, Min-Hsiu Shih, Yung-Nien Sun, and Ting-Li Shen. Pupil localization for ophthalmic diagnosis using anchor ellipse regression. In *2019 16th International Conference on Machine Vision Applications (MVA)*, pages 1–5. IEEE, 2019.

[21] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C Berg. Ssd: Single shot multibox detector. In *European Conference on Computer Vision*, pages 21–37. Springer, 2016.

[22] Changsheng Lu, Siyu Xia, Ming Shao, and Yun Fu. Arc-support line segments revisited: An efficient high-quality ellipse detection. *IEEE Transactions on Image Processing*, 29:768–781, 2019.

[23] Fei Mai, YS Hung, Huang Zhong, and WF Sze. A hierarchical approach for fast and robust ellipse extraction. *Pattern Recognition*, 41(8):2512–2524, 2008.

[24] HK Muammar and M Nixon. Tristage hough transform for multiple ellipse extraction. *IEE Proceedings E (Computers and Digital Techniques)*, 138(1):27–35, 1991.

[25] Ingram Olkin and Friedrich Pukelsheim. The distance between two random vectors with given dispersion matrices. *Linear Algebra and its Applications*, 48:257–263, 1982.

[26] Nobuyuki Otsu. A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1):62–66, 1979.

[27] Dilip K Prasad and Maylor KH Leung. An ellipse detection method for real images. In *2010 25th International Conference of Image and Vision Computing New Zealand*, pages 1–8. IEEE, 2010.

[28] Dilip K Prasad, Maylor KH Leung, Chai Quek, and Michael S Brown. Deb: Definite error bounded tangent estimator for digital curves. *IEEE Transactions on Image Processing*, 23(10):4297–4310, 2014.

[29] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 779–788, 2016.

[30] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems*, pages 91–99, 2015.

[31] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

[32] Christian Teutsch, Dirk Berndt, Erik Trostmann, and Michael Weber. Real-time detection of elliptic shapes for automated object recognition and object tracking. In *Machine Vision Applications in Industrial Inspection XIV*, volume 6070, page 60700J. International Society for Optics and Photonics, 2006.

[33] CL Todoroki, Eini C Lowell, and Dennis Dykstra. Automated knot detection with visual post-processing of douglas-fir veneer images. *Computers and Electronics in Agriculture*, 70(1):163–171, 2010.

[34] Veli-Matti O Tormanen and Anssi J Makynen. Detection of knots in veneer surface by using laser scattering based on the tracheid effect. In *2009 IEEE Instrumentation and Measurement Technology Conference*, pages 1439–1443. IEEE, 2009.

[35] S Tsuji and F Matsumoto. Detection of ellipses by a modified hough transformation. *IEEE Transactions on Computers*, 8(C-27):777–781, 1978.

[36] Gangyi Wang, Guanghui Ren, Zhilu Wu, Yaqin Zhao, and Lihui Jiang. A fast and robust ellipse-detection method based on sorted merging. *The Scientific World Journal*, 2014, 2014.

[37] Zeya Wang, Nanqing Dong, Sean D Rosario, Min Xu, Pengtao Xie, and Eric P Xing. Ellipse detection of optic disc-and-cup boundary in fundus images. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 601–604. IEEE, 2019.

[38] Jianhua Yang, Jiang Xiao, Wansi Fu, Jiawen Chen, Lei Yan, and Shuangyong Wang. Image processing and identification of lumber surface knots. *Acta Technica*, 62(1A):99–108, 2017.