IE 423 Quality Engineering Project Part 3

QUALITY CONTROL ON IMAGES

YAĞMUR ÖZDEMİR

UTKU BOYAR

İREM YAŞAR

Table of Contents

1. Introduction:

What is linen?

Linen, made from the fibres of the flax plant, distinguishes itself in the textile industry due to its exceptional characteristics. The fabric is highly esteemed for its exceptional durability, surpassing that of cotton, as well as its remarkable comfort. This organic fibre is renowned for its thermoregulatory properties, which make it a favored option across diverse sectors. Linen is utilized in various fields, including fashion, where it is highly valued for its sophisticated and neat look, as well as in hospitality and home decor, where its visual appeal and practical characteristics are highly regarded. The cultivation and processing of flax into linen is an ancient practice that showcases a combination of cultural legacy and skilled artistry. This enhances linen's standing in the market, establishing it as synonymous with excellence and opulence. [(1)]

Why is it important to monitor processing of linens?

Supervising the processing of linens is a crucial element in maintaining its exceptional standard of quality. Throughout the manufacturing process, the linen is at risk of a range of defects, including irregular weaving, discrepancies in color, and inconsistencies in texture. These imperfections can greatly affect the visual attractiveness and practical excellence of linen products, leading to a decrease in their market worth and consumer desirability. In sectors where linen represents opulence and coziness, such as upscale hotels or fashion, the existence of imperfections can significantly damage the brand's image. Hence, it is imperative to implement rigorous quality control measures, which involve meticulous monitoring and inspection throughout the processing stage. This guarantees that the end products satisfy the elevated standards of consumers and uphold the esteemed reputation of linen in the international market. [(1)]

The motivations of the use of images and identification of defects in linen manufacturing

The incorporation of image-based technology in linen manufacturing plays a crucial role in ensuring quality control. Manufacturers can significantly improve efficiency and accuracy in the production process by utilizing sophisticated image analysis techniques to identify defects. This approach not only enhances the overall excellence of linen goods but also optimizes manufacturing procedures, resulting in decreased wastage and heightened consumer contentment. The rationale for employing image analysis in the linen manufacturing industry arises from the imperative for accuracy and the substantial financial repercussions of flaws in this domain. [2]

2. Background information:

What has been done in the literature regarding the process monitoring on linen?

The process monitoring in linen manufacturing has gone through substantial advancements, combining conventional techniques with contemporary technological innovations to fulfill the increasing requirements for sustainability and effectiveness.

Historically, the labor-intensive process of linen production begins with the cultivation of flax plants. This encompasses the retting procedure, in which the flax stalks are immersed in water to decompose the pectin. Subsequently, the stalks are dried, threshed, scutched, and ultimately spun and woven into fabric. Although there have been improvements, a significant number of these procedures still rely on manual labor, necessitating skilled workers and meticulous attention to detail. This contributes to the perceived worth and artistry associated with linen. [3]

Traditional and sustainable methods of linen production have experienced a renewed interest in recent years. Artisans and small-scale producers are currently adopting organic farming and natural dyeing techniques, resulting in the production of eco-friendly fabrics. In addition, the industry has witnessed the implementation of automated weaving machines and digital imaging techniques for pattern generation, which has resulted in a more efficient manufacturing process and improved productivity. [4]

The fashion industry has experienced a resurgence in the use of linen due to its environmentally friendly attributes, including biodegradability, renewability, and its ability to utilize fewer resources compared to textiles such as cotton. Flax cultivation is renowned for its minimal water and pesticide needs, rendering it an ecologically sound alternative. The flax plant is fully utilized, resulting in a production process that generates no waste. [3]

In addition, the linen industry is currently exploring novel dyeing techniques that are ecologically sustainable, necessitating reduced water and energy consumption. Additionally, innovations encompass the creation of linen blends with alternative fibers, augmenting characteristics such as resistance to wrinkles and the ability to stretch. [4]

The integration of traditional craftsmanship and modern technological advancements in linen production not only upholds the fabric's historical and cultural importance but also aligns it with present-day sustainability objectives, rendering linen a viable and valuable textile in today's market.

3. Our Approach

Our methodology for detecting anomalies within textile images leveraged both statistical and machine learning techniques, particularly Principal Component Analysis (PCA) and k-Nearest Neighbors (KNN). These methods were selected for their distinct capabilities in handling high-dimensional data and their suitability for the task at hand, which involves the identification of defects or outliers in image data.

Principal Component Analysis (PCA)

PCA served as our primary tool for dimensionality reduction, allowing us to transform the image data into a lower-dimensional space that captures the most significant features. By focusing on the principal components that account for the highest variance within the dataset, we could effectively reconstruct the image while filtering out noise and minor inconsistencies. This reconstruction emphasized the key structural elements of the image, thereby enabling us to identify anomalies as deviations from the reconstructed norm. The PCA model was applied to patches extracted from the images, which were then reassembled to visualize the outliers.

k-Nearest Neighbors (KNN)

Concurrently, we utilized the KNN algorithm to predict each pixel's intensity based on the surrounding neighbors. This localized approach aimed to maintain the textural integrity of the image while smoothing out irregularities. We used grid search to optimize the 'k' parameter, ensuring that the model neither overfit nor underfit the data. The optimal 'k' value was determined through cross-validation, and the resulting reconstructed images were compared against the original to calculate the residuals. These residuals were then analyzed using control chart methods to identify outlier pixels.

Image Analysis

For both PCA and KNN methods, we proceeded by calculating the residuals between the original and reconstructed images. These residuals were scrutinized for outliers using control limits established through statistical thresholds (e.g., 3-sigma or 4-sigma rules). Outliers identified through this process were indicative of potential defects or anomalies in the textile material.

Considerations in Our Approach

Throughout our analysis, we were cognizant of the inherent characteristics of each method. PCA's global perspective and KNN's local approach were both considered to offer complementary insights into the data. The comparative analysis of the two methods allowed us to understand their respective strengths and limitations in the context of anomaly detection within complex image data.

In our application of these methodologies, we navigated challenges such as the appropriate choice of hyperparameters, computational efficiency, and the balance between sensitivity to anomalies and robustness against noise. The versatility of PCA in capturing the most impactful elements of an image, coupled with the nuanced local analysis provided by KNN, formed the bedrock of our approach to identifying outliers in textile images.

4. Results

Our application of PCA and KNN models on a set of textile images yielded insightful results, demonstrating the effectiveness of each method in detecting potential anomalies within the images.

PCA Results:

The PCA model, when applied to image patches, was able to distill the essence of the images by focusing on the components with the highest variance. The reconstructed images, though smoother, preserved the primary patterns and textures, albeit with a reduction in finer details.

This allowed for the more prominent anomalies to stand out in the residual analysis. The "Outliers (4-sigma)" images, highlighted by the PCA method, presented scattered points across the images, pinpointing the significant deviations from the reconstructed image. These anomalies correspond to pixels that vary substantially from the primary components, suggesting potential defects.

KNN Results:

The KNN model provided a localized analysis, where each pixel's intensity was predicted based on its nearest neighbors. This resulted in a smoothing effect, which was more pronounced in the reconstructed images, sometimes obscuring subtle defects. However, the KNN method's sensitivity to local variations led to a denser distribution of potential outliers across the images. The "Outliers (3-sigma)" visualization showed numerous points of interest, suggesting that KNN might be overfitting to the texture noise and possibly generating false positives.

Comparison and Discussion:

In our comparative analysis of Principal Component Analysis (PCA) and k-Nearest Neighbors (KNN) for outlier detection in textile images, PCA demonstrated a more effective performance for several reasons:

1. Dimensionality Reduction and Feature Capturing: PCA's strength lies in its dimensionality reduction capability, which is crucial for capturing the most variance in the data. By identifying principal components that encapsulate significant features of the image, such as edges, textures, and patterns, PCA can effectively reconstruct the key elements of an image. This focused reconstruction allows PCA to disregard noise and highlight more pronounced outliers compared to the original image.

2. Localized vs Global Analysis: The PCA approach contrasts with the localized analysis of KNN, which predicts each pixel's value based on its nearest neighbors. While KNN can preserve local similarities, it tends to over-smooth the image, potentially masking subtle anomalies. PCA, however, analyzes the global structure of the image, offering a holistic view that is often more effective in identifying areas that deviate significantly from the norm.

3. Computational Efficiency: In terms of computational demands, KNN, which calculates distances between pixels, can be resource-intensive, particularly for high-resolution images. PCA, conversely, involves linear algebra operations on the data's covariance matrix. Once the principal components are determined, reconstructing the image or pinpointing outliers becomes more computationally efficient in PCA compared to the distance-based approach of KNN.

4. Robustness to Variance and Consistency in Results: PCA's focus on high-variance components of an image makes it robust in identifying areas that deviate markedly from the average, often corresponding to defects or anomalies. This robustness is reflected in the consistency of the explained variance across different images, which hovered around 40%. This level of variance indicates PCA's ability to capture a substantial amount of information in the textile images, providing a reliable baseline for comparison and anomaly detection.

In summary, PCA's capability to distill and reconstruct images based on their most defining characteristics, coupled with its computational efficiency and robustness in highlighting significant deviations, renders it particularly suitable for tasks like quality control in image analysis. The consistent performance across different images, with a similar variance explained, reinforces PCA's suitability for identifying unusual or anomalous features in such applications.

5. Conclusions and Future Work

Our exploration into the use of PCA and KNN for detecting anomalies in textile images led to several conclusions. PCA's strength lies in its ability to abstract and reconstruct images based on capturing the most variance, thus effectively highlighting significant deviations when anomalies are present. In contrast, KNN's localized approach, while valuable for preserving textural information, may be less suited to this application due to its tendency to smooth over subtle defects and potentially overfit to noise.

Conclusions

- PCA provided a more streamlined and computationally efficient method for identifying potential defects in the images, benefiting from its dimensionality reduction capabilities and focus on the most significant features within the image data.

- The KNN approach, despite its optimal k value determined via grid search, exhibited a propensity for overfitting, resulting in a high density of outliers that included many false positives.
- The controlled environment of the analysis, which lacks external validation, highlights the need for further testing and refinement of these methods to ensure robustness and accuracy in real-world settings.

Future Work

To build on the foundation laid by our initial approaches, future work could include:

- Hybrid Modeling: Combining PCA with machine learning algorithms like clustering or with other dimensionality reduction techniques could lead to a more nuanced detection of anomalies.
- Deep Learning Integration: Leveraging convolutional neural networks (CNNs) and autoencoders could enhance feature extraction and anomaly detection capabilities, especially in capturing complex patterns.
- Ensemble Methods: Using an ensemble of various models may improve performance by balancing the strengths and weaknesses of individual approaches.
- Cross-validation with Expert Annotation: Incorporating expert input and cross-validation with labeled data can help in fine-tuning the models and validating the anomalies detected.
- Adjusting Hyperparameters:** For KNN, experimenting with different values of k and other parameters may yield a more accurate detection of true outliers.
- Expanding the Dataset: Applying the PCA and KNN models to a larger and more diverse set of images would test the scalability and generalizability of the findings.
- Noise Reduction and Feature Engineering: Pre-processing images with noise reduction techniques and engineering features that could better represent the defects could improve the sensitivity and specificity of the models.
- Automated Threshold Determination: Developing methods for automatic threshold setting in outlier detection could make the process more adaptive to different types of images and anomalies.

These extensions could greatly enhance the robustness of the detection approach, ultimately leading to more reliable quality control in the textile industry and beyond.

References

1) https://www.wilkinslinen.com/textiles/linen-production-process-made/
2) Manufacturing processes in the textile industry. Expert Systems for fabrics production. Bullon, Juan & González, Angélica & Encinas, Ascensión & Queiruga-Dios, Araceli. (2017).
3) https://cosh.eco/en/articles/climate-resilient-and-versatile-linen-the-future-proof-fabric-for-fashion
4) https://nancysnotions.com/the-future-of-linen-fabric-innovations-and-sustainable-practices/

ChatGPT Prompts

For Principal Component Analysis:

"residual image is an empty white box, what is the reason behind this problem and how can we fix it?" "even using the new code and pca(0.5) does not result in any residuals and it is still appearing empty, you said this: Nature of the Image: If the image is very uniform (little variation in pixel values), even a small number of PCA components might reconstruct it very accurately.

Solution: If the nature of the image is such, you might need to consider a different approach for defect detection, or focus on smaller regions of the image. how can i modify the code so that it will focus on smaller regions of the image"

For KNN Approach:

"okay this is great, now we need to generate another approach for this outlier detection. I want you to write a python code that will use k-nearest neighbors approach (you should find the optimum k for our image) then it should predict each pixel with the knn approach"

For overall comparison part: "now these are our alternative image names: #0089.jpg, 0097.jpg, 0170.jpg, 0173.jpg and 0192.jpg we need to test them on our pca approach which was this code:

--our existing codes--

can you turn this code into the related python function for testing these new images"

"now we need to test our images on our knn approach which was this code:

--our existing codes--

can you turn this code into the related python function for testing these new images"