POLITECNICO
MILANO 1863

SCUOLA DI INGEGNERIA INDUSTRIALE
E DELL'INFORMAZIONE

Executive Summary of the Thesis

Multispectral imaging for solvent classification using NIR & Visible spectra

Laurea Magistrale in Mathematical Engineering - Ingegneria Matematica

Author: Martino Riccardi

Advisor: Prof. Marco Marcon

Co-advisor: Daniele Ugo Leonzio

Academic year: 2022-2023

1. Introduction

The executive summary provides an overview of the study conducted on multispectral imaging and the utilization of visible and near-infrared (NIR) images for solvent classification [1]. The study focuses on a two-terminal dual-band detector that captures images of different solvents by switching the bias polarity between the near-infrared and visible bands [2]. The objective is to develop classification models that leverage both the visible and NIR information simultaneously and determine if this approach improves performance compared to using them separately. The study explores the computation of the ratio between the IR and visible components to mitigate illumination effects and prevent classifiers from detecting offsets.

The main steps of the study involve acquiring a checkerboard image for optimal illumination and focus conditions and testing focus measures. After finding the best distance, a dataset of multispectral images of liquid bottles has been acquired, which had to undergo some preprocessing stages like addressing vignetting effects, extracting features from portions of liquid and applying PCA for dimensionality reduction. Finally the dataset could be used for classification purposes using various classifiers and evaluating their performance.

The results of the study demonstrate the effectiveness of the developed models in accurately identifying and differentiating solvents. The performance of the classifiers is extensively discussed and the use of the combined visible and NIR information is found to enhance the classification accuracy. Additional images are used for predictions, validating the robustness and generalizability of the obtained results.

2. Data Acquisition

In this section, we provide an overview of the data acquisition process and focus evaluation conducted in our study. Our sensor captured both visible and near-infrared (NIR) images at various distances to determine the optimal focus distance. To assess focus quality, we utilized a checkerboard pattern that offered well-defined features, spatial frequency analysis and calibration/reference capabilities.

To ensure accurate focus evaluation, we implemented preprocessing steps to address vignetting effects caused by uneven illumination across the images. By cropping the central portion of the checkerboard image and applying

grayscale normalization, we improved the accuracy of focus measures. Additionally, rescaling the images using global maximum and minimum values allowed for fair comparison and eliminated potential biases.
We explored different types of focus measures, including spatiality-based, statistics-based and frequency-based methods. After careful evaluation, we selected the Laplacian (LAPD) approach for its ability to capture sign changes and detect finer details. We compared the scores of focus measures at different distances, which enabled us to determine the optimal focus distance for both silicon-based and germanium-based photodiodes.
Based on our analysis, a focal distance of 22.9433mm had been suggested, which is obtained by averaging the recommended distances for silicon and germanium photodiodes. By employing this determined distance for image acquisition, we ensure that the captured images are in focus and suitable for subsequent processing and analysis. This approach significantly enhances the accuracy and reliability of our imaging system, leading to improved results in our study.

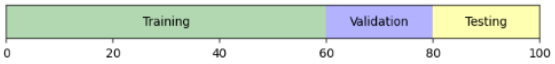
3. Preprocessing

Prior to initiating the classification procedure, it is essential to perform preprocessing steps to ensure the availability of informative and undistorted images. These stages primarily involve correcting image aberrations, extracting relevant image regions containing the liquid and combining visible (VIS) and near-infrared (NIR) information [3]. Two distinct approaches were employed for image extraction: the meniscus-encompassing portion and the internal liquid region. The former captures a region that includes the background and a segment of the liquid, while the latter focuses solely on the liquid within the bottle. These approaches yield different dataset sizes and provide flexibility in analysis. For the purpose of this work, we will concentrate on the internal liquid approach, as it has yielded the most favorable and informative results. The analysis begins by capturing images of the bottles, both in the visible and infrared spectra. To address vignetting, which can lead to uneven illumination, the images are corrected by multiplying the inverse blank matrix. Some portions, entirely contained within the bottles' perimeters, have been extracted and, alongside utilizing singularly NIR and Visibile information, those portions have been used to create two new datasets: in one case the ratio of NIR to VIS pixels has been calculated. Alternatively, the NIR and VIS data can be kept as separate vectors, but this approach carries the risk of capturing differences unrelated to the intrinsic nature of the liquids but related to the sensor characteristics. Subsequently, 3x3 squares are extracted from all the portions: these squares are then flattened and the obtained vectors become the rows of the data matrices which serve as the basis for the classification process. Eventually PCA is employed for dimensionality reduction of the data matrices and to determine the optimal number of principal components for each dataset, with the cumulative sum of eigenvalues serving as a guide. The analysis reveals that, in most cases, a single principal component is sufficient to explain the variance in the dataset, considering only the liquid properties. However two principal components are utilized.
In conclusion, by meticulously executing these preprocessing steps, the dataset is suitably prepared for subsequent classification and analysis tasks. This comprehensive approach facilitates the precise differentiation of diverse solvents based on the captured images, providing valuable insights into their unique spectral properties.

4. Classification theory

This section delves into the theoretical aspects of supervised learning techniques for classification tasks. Supervised learning involves training algorithms using labeled datasets to accurately classify or predict outcomes. A crucial aspect in model development is finding the right balance between underfitting and overfitting. To achieve this, two common validation techniques are used.
The first technique is hold-out validation, where the dataset is split into training, validation, and test subsets. This approach is computationally efficient but may result in biased estimates and reduces the amount of data available for training. The second technique is cross-validation, which involves using multiple subsets or folds for training and validation. This provides more ro-

bust performance estimates but increases computational complexity. For the internal liquid approach, hold-out validation will be adopted with a splitting of 60% training set, 20% validation set and 20% testing set [4].

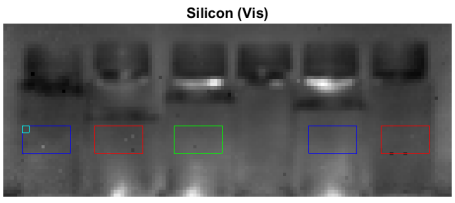


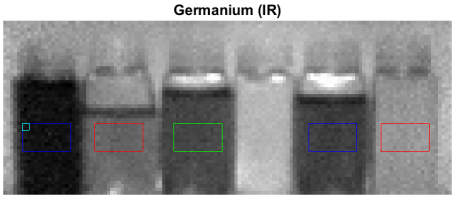
The algorithms used in this work are discriminant classifiers, Support Vector Machines (SVM), decision trees and KNN. These classifiers are trained using the labeled training set and their hyperparameters are optimized [5].
Discriminant classifiers include Linear Discriminant Analysis (LDA) and Quadratic Discriminant Analysis (QDA). LDA assumes a linear combination of features and aims to identify the discriminant function that maximally separates classes. QDA relaxes the linearity assumption, allowing for nonlinear classification decision boundaries.
SVM find optimal hyperplanes that separate classes. Kernel functions are used to handle non-linear decision boundaries. Different strategies, such as One-vs-One and One-vs-All, can be adopted for multiclass problems. Hyperparameter tuning is crucial for SVMs and involves selecting appropriate values for parameters such as the choice of kernel function, the regularization parameter (C), and the kernel coefficient (gamma) for specific kernel functions.
K-nearest neighbor (KNN) algorithm finds the K closest data points to a query point based on a chosen distance function and uses their labels for prediction. KNN is a non-parametric method that does not assume any underlying data distribution. Different distance measures, such as Euclidean and Manhattan, can be used to calculate the similarity or dissimilarity between data points. The choice of the value of K in KNN is important, as it determines the influence of neighboring points on the classification decision ("fine KNN" if k=1, "medium KNN" if k=10).
Decision trees involve recursively splitting the data based on attribute values. At each decision node, a test function is applied to determine the appropriate branch to follow and this process continues until a leaf node is reached, providing the final output. Decision trees are non-parametric models and do not rely on specific assumptions about class densities. Various impurity measures, such as entropy, are utilized to assess the quality of splits in classification trees. By examining accuracy values and confusion matrices, informed decisions can be made regarding the selection of the most suitable classifier for a given task. Comparing the results of different classifiers helps identify those demonstrating better overall performance.

5. Classification results

The first analysis involves images of five bottles, filled with the following liquids: water, acetone, isopropanol, ethanol and an empty vial. There is actually an additional bottle containing Toluene but the content evaporated during the acquisition so it has been discarded. The second analysis involves images of six bottles, filled with: water, isopropanol, water+iso, an empty vial, ethanol and acetone. After carefully following the steps of the preprocessing described in previous session (with Internal Liquid approach), the classification process could begin. Every matrix data could underwent the assignment of class labels based on the liquid properties and we trained the classifiers with a splitting of 60%, 20%, 20%. The evaluation of the classification models shows promising results, with the accuracy of the classification process reaching high levels.

5.1. 5 liquids analysis





Starting from the NIR analysis, looking at the table, the accuracies are all high, but none of the classifiers reaches the 100% accuracy as the Isopropanol and Ethanol proved to be difficult to distinguish. QDA doesn't improve LDA results, a Gaussian kernel has been used for the SVM,

there's no particular difference between fine and medium decision tree and Manhattan distance with k=10 proves to be optimal for KNN.

Method	Validation	Testing
LDA	0.923	0.949
SVM	0.949	0.974
Tree	0.949	0.974
KNN	0.923	0.923

Considering the visible data, the accuracy scores indicate that the classifiers struggle to accurately assign labels to the data points, confirming the complex nature of the data. The best accuracy is provided by SVM (with Gaussian Kernel)

Method	Validation	Testing
QDA	0.641	0.615
SVM	0.667	0.744
Tree	0.590	0.615
KNN	0.513	0.667

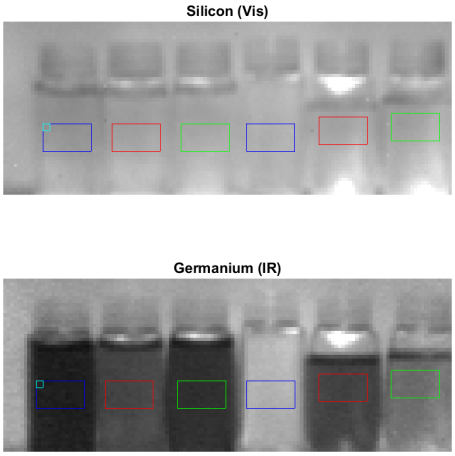
Continuing with the ratio approach, in contrast to the sole use of NIR (Near-Infrared) data, this analysis incorporates classifiers that demonstrate accuracies of 100%. This improvement in classification accuracy might indicate that using ratioed information has a positive impact on the classification process. Here the QDA seems to improve the LDA results, a Gaussian kernel has been used for SVM, fine and medium trees are both effective and for KNN we used a weighted Manhattan distance (inverse) and k=10.

Method	Validation	Testing
QDA	0.974	1.0
SVM	0.949	1.0
Tree	0.949	0.974
KNN	0.949	0.949

Finally, the rows of the visible matrix were appended to the infrared matrix, allowing the preservation of both pieces of information separately. The accuracy are still high but we stress that in this case we might risk to detect offsets caused by the sensor, not the liquids. The hyperparameters are basically the same as the ratio approach: for KNN the distance is unweighted and QDA doesn't improve LDA results.

Method	Validation	Testing
LDA	0.974	1.0
SVM	0.974	1.0
Tree	0.974	0.974
KNN	1.0	0.974

5.2. 6 liquids analysis



Starting again with the NIR analysis, the accuracies are still high, however with worse results than before (the underlying distinction between Isopropyl Alcohol and Ethanol appears more challenging).

Method	Validation	Testing
QDA	0.915	0.913
SVM	0.915	0.935
Tree	0.851	0.913
KNN	0.894	0.891

Continuing with the visible data, the obtained accuracy scores indicate that the classifiers face difficulties in accurately assigning labels to the data points, providing evidence of the intricate and perplexing nature of the data.

Method	Validation	Testing
LDA	0.660	0.587
SVM	0.553	0.565
Tree	0.413	0.413
KNN	0.447	0.500

Considering the ratio approach, our analysis incorporates classifiers that demonstrate accuracies close to 100%. Again, this improvement in classification accuracy might indicate that using ratioed information has a positive impact on the classification process.

Method	Validation	Testing
LDA	0.957	0.978
SVM	0.957	0.978
Tree	0.957	0.935
KNN	0.957	0.913

At last, combining the visible and infrared matrices by appending the rows, the classifiers show relatively high accuracies, but not as much as the ratio approach

Method	Validation	Testing
QDA	0.872	0.913
SVM	0.915	0.978
Tree	0.936	0.935
KNN	0.915	0.935

5.3. Final considerations

Using the ratio of NIR to Vis proves to be more effective, harnessing the complementary nature of the two types of data. This ratio-based approach shows notable improvements in accuracy across different classifiers, especially in the 6-liquid analysis. The ratio-based approach has shown high predictive capabilities and these capabilities will be further evaluated in the final chapter using small sections of liquids from different images. The goal is to predict the majority of vectors within each section and determine a dominant class label, effectively characterizing the overall composition of the liquid section.

6. Prediction results

After training the classifiers, the models can be used for prediction on a new set of images taken at a different time. The new images contain portions of liquids that need to be classified undergoing the same previous preprocessing steps. Evaluation metrics, in addition to accuracy, are used to assess the performance of the classifiers on the new data. These metrics provide quantitative measures of accuracy, precision, recall and F1 score [6]. Specificity measures the ability to correctly identify negative instances, precision assesses the accuracy of positive predictions, recall evaluates the ability to identify positive instances and the F1 score provides a balanced measure of precision and recall. Recall in particular is useful to determine a dominant class for any section, so even if we don't classify all the vectors correctly, we can classify the liquid section the right way. The results are shown below:

	QDA	SVM	Tree	KNN
Accuracy	0.9538	0.9538	0.9436	0.9487

The discriminant and the SVM classifiers show better performances, even though all the accuracies are all pretty high.
Focusing on the evaluation metrics:

QDA	H_2O	Ace	Iso	Etha	Empty
Specificity	1.0	1.0	0.97	0.97	1.0
Recall	1.0	1.0	0.90	0.87	1.0
Precision	1.0	1.0	0.88	0.90	1.0
F1 Score	1.0	1.0	0.89	0.88	1.0

SVM	H_2O	Ace	Iso	Etha	Empty
Specificity	1.0	1.0	0.97	0.97	1.0
Recall	1.0	1.0	0.87	0.90	1.0
Precision	1.0	1.0	0.90	0.88	1.0
F1 Score	1.0	1.0	0.88	0.89	1.0

Tree	H_2O	Ace	Iso	Etha	Empty
Specificity	1.0	1.0	0.98	0.95	1.0
Recall	1.0	1.0	0.79	0.92	1.0
Precision	1.0	1.0	0.91	0.82	1.0
F1 Score	1.0	1.0	0.85	0.87	1.0

KNN	H_2O	Ace	Iso	Etha	Empty
Specificity	1.0	1.0	0.98	0.96	1.0
Recall	1.0	1.0	0.82	0.92	1.0
Precision	1.0	1.0	0.91	0.84	1.0
F1 Score	1.0	1.0	0.86	0.88	1.0

The evaluation of the classifiers on new, unseen images confirms their strong performance in classifying the liquids, validating the results from Section 5. As explained before the most problematic solvents are the Isopropanol and the

Ethanol, confirmed by the evaluation metrics. The classifiers have successfully learned the distinguishing features and patterns of the liquids, enabling them to generalize well. Among the four classifiers, QDA and SVM exhibit slightly better overall accuracy and precision compared to Tree and KNN. However, Tree and KNN classifiers still provide reliable results, albeit with slightly lower scores in some metrics.

7. Conclusion

This thesis focused on applying machine learning techniques to classify multiple solvents using visible and near-infrared (NIR) images. The NIR over Visible approach was found to be the most effective strategy, achieving high accuracy rates exceeding 95% across various classifiers, with discriminant analysis and support vector machines performing particularly well. Decision trees and k-nearest neighbors algorithms also produced satisfactory results. Independent dataset predictions confirmed the robustness of the classification procedure, consistently achieving accuracies around 95%.

The computational efficiency of the training and prediction processes suggests potential for further improvements. However, certain limitations were identified, including the impact of lighting conditions on analysis accuracy and the importance of acquiring high-quality images with optimal focus. The limited availability of data due to the time-intensive acquisition process and the need for exploring additional features beyond light information were also acknowledged.

Future developments could involve expanding the dataset to encompass a wider range of solvents, including those with varying transparency in the visible spectrum. Exploring different types of oils, where NIR information may not be as discriminatory, could present a challenging task. Additionally the exploration of the Short-Wave Infrared (SWIR) spectrum range could provide improvements for capturing subtle differences. Incorporating spatial information into the analysis and utilizing more advanced classifiers such as ensemble techniques or neural networks could further enhance classification accuracy. These avenues offer opportunities for future research and improvements in the field.

8. Acknowledgements

I express my gratitude to my advisor Professor Marco Marcon, my co-advisor Diego Ugo Leonzio, Professor Giovanni Isella, Andrea Ballabio and Professor Jacopo Frigerio for their guidance and support throughout the research process of this work.

References

[1] E Talamas Simola, A De Iacovo, J Frigerio, A Ballabio, A Fabbri, G Isella, and L Colace. Voltage-tunable dual-band ge/si photodetector operating in vis and nir spectral range. *Optics express*, 27(6):8529–8539, 2019.

[2] Enrico Talamas Simola, Vivien Kiyek, Andrea Ballabio, Viktoria Schlykow, Jacopo Frigerio, Carlo Zucchetti, Andrea De Iacovo, Lorenzo Colace, Yuji Yamamoto, Giovanni Capellini, et al. Cmos-compatible bias-tunable dual-band detector based on gesn/ge/si coupled photodiodes. *ACS photonics*, 8(7):2166–2173, 2021.

[3] Yuanjie Zheng, Stephen Lin, Chandra Kambhamettu, Jingyi Yu, and Sing Bing Kang. Single-image vignetting correction. *IEEE transactions on pattern analysis and machine intelligence*, 31(12):2243–2256, 2008.

[4] Ajitesh Kumar. Hold-out method for training machine learning models. *Data Analytics*, 2021.

[5] Ethem Alpaydin. *Introduction to machine learning*. MIT press, 2020.

[6] Mohammad Hossin and Md Nasir Sulaiman. A review on evaluation metrics for data classification evaluations. *International journal of data mining & knowledge management process*, 5(2):1, 2015.