Learning and Nonlinear Models - Journal of the Brazilian Society on Computational Intelligence (SBIC), Vol. 21, Iss. 2, pp. 16-28, 2023

ANALYSIS AND IMPROVEMENT OF MACHINE LEARNING MODELS FOR DETECTING STREET LIGHTING LAMPS

Igor Caetano Silva, **Ricardo Menezes Salgado**
Institute of Exact Sciences – Federal University of Alfenas
igor.caetano@sou.unifal-mg.edu.br, ricardo.salgado@unifal-mg.edu.br

Igor Mattos dos Santos Varejão, **Flavio Miguel Varejão**
Department of Informatics – Federal University of Espírito Santo
fvarejao@inf.ufes.br, igor.varejao@edu.ufes.br

Resumo – A rede de iluminação pública brasileira é mantida pelas prefeituras. Para o faturamento da energia fornecida às prefeituras, as distribuidoras de energia deverão manter um banco de dados atualizado dos postes da rede e seus tipos de lâmpadas e potências. Porém, é comum haver problemas com desinformação, em que a empresa não é notificada sobre alterações na rede de iluminação pública pelas prefeituras e não consegue atualizar adequadamente seu banco de dados. Para evitar perdas comerciais, as empresas enviam equipes para verificar manualmente a infraestrutura, um processo caro, demorado e pouco confiável. Nesse sentido, este trabalho visa otimizar os modelos propostos na literatura capazes de classificar com precisão o tipo e a potência das lâmpadas em postes de iluminação pública com base em dados coletados de sensores radiométricos e de uma câmera profissional. Dessa forma, os dados são processados usando algoritmos tradicionais de aprendizado de máquina e aprendizado profundo, juntamente com técnicas de validação mais complexas de transformação de dados e otimização de hiperparâmetros para obter melhores resultados. Com base nesta metodologia, os resultados mostram que modelos com algoritmos mais robustos (*Support Vector Machine*, *XGBoost*, *Random Forest* e *Multilayer Perceptron*) conseguem atingir uma acurácia média final de 80-86%, o que confirma a utilidade desta metodologia como uma solução alternativa para reduzir as perdas comerciais da iluminação pública.

Palavras-chave – Aprendizado de máquina, Iluminação de rua, Iluminação pública, Processamento de imagens, Sensores de luz.

Abstract – The Brazilian public lighting network is maintained by city halls. To bill the energy provided to city halls, energy distribution companies should maintain an updated database of network poles, their lamp types, and wattages. However, it is common to encounter issues with misinformation, where the company is not notified about changes in the public lighting network by city halls and cannot update its database appropriately. To mitigate commercial losses, companies have resorted to sending teams for manual infrastructure checks, which is an expensive, time-consuming, and unreliable process. In this regard, this work aims to optimize the models proposed in the literature capable of accurately classifying the type and wattage of lamps on public lighting poles based on data collected from radiometric sensors and a professional camera. Data is processed using traditional machine learning and deep learning algorithms, along with more sophisticated validation techniques such as data transformation and hyperparameter optimization to achieve improved results. Based on this methodology, the results demonstrate that models employing more robust algorithms (Support Vector Machine, XGBoost, Random Forest, and Multilayer Perceptron) can attain a final average accuracy of 80-86%. This confirms the usefulness of this methodology as an alternative solution to address the issue of public lighting billing.

Keywords – Image processing, Light sensors, Machine learning, Public lighting, Street lighting

1. INTRODUCTION

Public lighting is defined as a public service whose exclusive purpose is to provide light (continuous, periodic, or occasional) to public places at night or during occasional daytime dimming. This service covers energy supply to the most diverse public places, such as streets, avenues, roads, squares, parks, public transport points, and historical or cultural monuments. Thus, it is an essential service for modern life (especially in urban centers) and is directly linked to public safety. It allows citizens to enjoy public spaces at night and guarantees traffic safety, helping encourages tourism and local commerce, among several other benefits.

In Brazil, the sole responsibility for the installation and maintenance of public lighting rested with the power distribution companies until 2010. However, due to changes in governmental regulations, these responsibilities were delegated to the municipalities. Nevertheless, the distribution companies still bear the responsibility for billing the city halls. As the city halls began to modify the lighting network and replace old luminaires, they often failed to accurately inform the power distribution companies about these changes. When alterations to the lighting infrastructure occur without proper documentation, the data on electricity consumption remains non-updated. This situation can lead to incorrect billing, where consumption is inaccurately attributed to different regions or entities. Consequently, this results in revenue losses for both municipalities and distribution companies, with broader economic implications. Therefore, it is a matter of great concern for regulators and electricity companies, as they may face significant economic, technical, and efficiency losses, as reported in [1].

To address this problem, companies began sending technical teams to inspect individual lighting points in the field and verify if the public lighting network matches their database. However, this process is slow and costly, often requiring inspectors to climb the pole to identify the correct lamp type and wattage, which can be of dubious effectiveness. Additionally, this solution can be unreliable, as it depends on the qualifications of the inspection team, whether internal or outsourced.

Another problem that arises is that, in some cases, technicians may not thoroughly inspect a lighting point and provide an approximate answer, leading to incorrect data or even omitting new lighting points from inspection. Public lighting networks can also feature numerous luminaire designs, varying lamp wattages, and a diverse range of public road types, which further complicates the task of identifying installations, making it both challenging and complex.

These conditions motivated the development of a device that incorporates both hardware elements (such as light sensors, a digital camera, and peripheral circuits) and software (computational techniques), with a high degree of automation. This device is designed to determine the type and wattage of lamps installed on poles used for public lighting. Soares et al. [2–4] have described the main elements of this device, how it operates, and reported test results based on laboratory and field-collected data. These promising results suggest a rapid update of public lighting records, potentially reducing the commercial losses resulting from the lack of energy consumption information. The device capable of acquiring ambient lighting information is installed on the roof of a vehicle (Figure 1). The collected data is processed by an intelligent algorithm that provides information about the location, type, and power of the lamps. Previous work [2–4] yielded relevant results but only a few algorithms and simpler validation methods were used in the experiments. The simpler validation methods may produce over optimistic results. Additionally, recent machine learning classification algorithms may offer improved performance compared to those used in previous work.

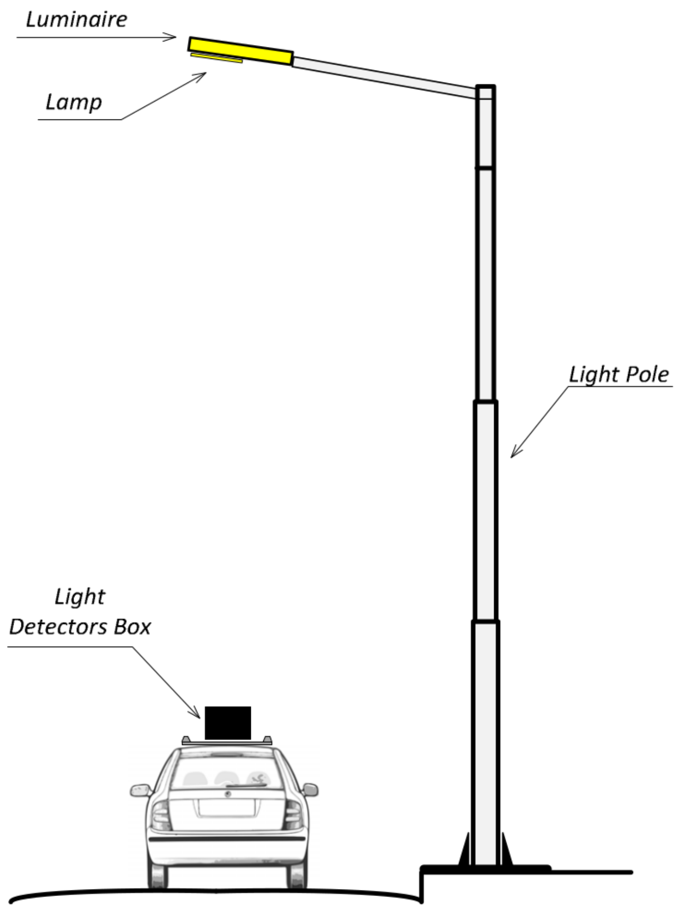


Figure 1: Vertical cut view of a typical acquisition scene including transportation vehicle, device case and street lighting equipment (extracted from [2]).

Hence, this work explores an enhanced version of prior research by employing a wider array of intelligent algorithms (of different approaches and levels of complexity) and a better validation method in order to achieve better and more realistic accuracy estimation in classifying lamp types and wattages. Experimental results have demonstrated that the more robust methods outperform those utilized in previous studies, affirming the generalizability of this approach for practical applications.

This paper is organized as follows: Section 2 reviews related work and points out some aspects that should be improved.

Section 3 briefly describes the eletronic device and the process of collecting data. Section 4 presents the theoretical background required for understanding this work, focusing on the image processing methods and the algorithms utilized to build the intelligent models. Section 5 describes the features of the dataset and the training and validation methodology used in the experiments. Section 6 presents the experimental results, our analysis and findings. Finally, Section 7 shows the conclusions and perspectives for future works.

2. RELATED WORKS

Public Lighting (PL) systems have recently become a focal point of research. Dwiyaniti et al. [5] address the need for modernizing public streetlight control and monitoring systems issues by developing a smart streetlight monitoring system using IoT (Internet of Things) technology. Kanthi and Dilli [6] directed their efforts toward minimizing power consumption within smart street lighting implementations while ensuring security through encrypted brightness level adjustments via a mobile application. Liu et al. [7] combines remote sensing technology with ground investigation, encompassing physical measurements of lighting attributes and surveys of pedestrians' safety and visual comfort perceptions, with multiple regression analysis, in order to evaluate the perceived quality of street lighting. Mavromatis et al. [8] utilized a substantial image dataset to monitor the proper functioning of PL lamps, ensuring their response aligns with illumination requirements. It's noteworthy that, while contributing to the broader field of public lighting, these studies do not directly address the specific challenge of classifying public lamps.

Part of the research presented in this work has been previously discussed in Soares et al. [2–4], where the system's hardware and software are detailed, as well as in Broetto et al. [9], which analyzed the application of image processing techniques to automatically detect street lighting luminaries, and Broetto and Varejão [10], which combines feature extraction and selection algorithms to optimize system classification accuracy.

Soares et al. [2] were the pioneers in proposing the computational methodology for extracting information about public lighting points. They utilized a dataset composed of radiometric sensor data captured in a laboratory in the classification experiments. In their preliminary experiments, they employed machine learning methods including K nearest neighbors [11], decision tree [12] and multilayer perceptron [13] along with a train-test split resampling method to evaluate the classifiers.

Broetto et al. [9] only used a dataset consisting of Fourier [14], Hu [15] and Haralick [16] descriptors of images of two types of luminaires. The images were obtained in laboratory and in field. They used two digital cameras with different resolution and zoom ratio for collecting the images. They also used K nearest neighbors, decision tree and multilayer perceptron as the machine learning techniques for classifying the images. The authors did not specify which validation method was used for gathering the experimental results.

Soares et al. [3] used a dataset obtained in a laboratory, simulating various environmental conditions. Radiometric and optical sensors in addition to the pole height compound the features of the dataset. An hierarchical architecture in two classification levels was employed. In the first level, the classifier identifies the lamp type and the second level classifier identifies the lamp wattage. For each classification level, the authors evaluated machine learning techniques, including K nearest neighbors, decision trees, and multilayer perceptrons. They used the stratified 10-fold cross-validation method in their experiments.

Soares et al. [4] used the same dataset, the same classifiers and the same hierarchical classifier architecture as in [3], but included a feature selection preprocessing step in order to select the best feature subset for improving the overall classification accuracy. They used stratified 10 fold cross-validation method for feature selection and the leave one out method for comparing the classifiers accuracy.

Broetto et al. [10] integrated data from the radiometric sensors and the image descriptors, along with RGB and EXIF descriptors obtained by the digital camera to form a heterogeneous pool of features. They only used actual (non laboratory) data collected in the field. Then, they applied three different feature selection wrapper methods selecting the best feature subset for improving the classification performance. They used stratified 10 fold cross validation and just the K nearest neighbors classifier in the experiments.

It's worth noting that each of the previous work basically uses different types of data on their datasets. They also always use the same traditional classifiers: k nearest neighbors, decision tree and multilayer perceptron, except for [10], which exclusively used k-nearest neighbors. Furthermore, these previous works typically used either a train-test split or stratified cross validation as resampling methods in the experiments. Both of these methods may potentially introduce bias into the hyperparameter tuning of the classifiers, which could lead to overly optimistic results.

In contrast to the previous approaches, this work uses a field collected dataset only composed of the radiometric sensors data and the image descriptor data. In addition to the traditional classifiers used in the previous approaches, this work includes robust new classifiers methods, such as, support vector machine, random forest and extreme gradient boosting. More importantly, this work uses a nested cross validation resampling method. This method prevents hyperparameter tuning from being biased by the testing dataset, ensuring that the results are less likely to be optimistic.

3. DATA ACQUISITION EQUIPMENT

The system proposed in [2–4] was designed to enhance the efficiency and cost-effectiveness of the inspection process. Therefore, the hardware that constitutes the overall system consists basically of low-cost and off-the-shelf radiometric sensors and a camera capable of collecting various types of information about the lighting equipment under evaluation. Figure 2 illustrates the equipment architecture, which was developed to serve as the main database to feed the computational system.

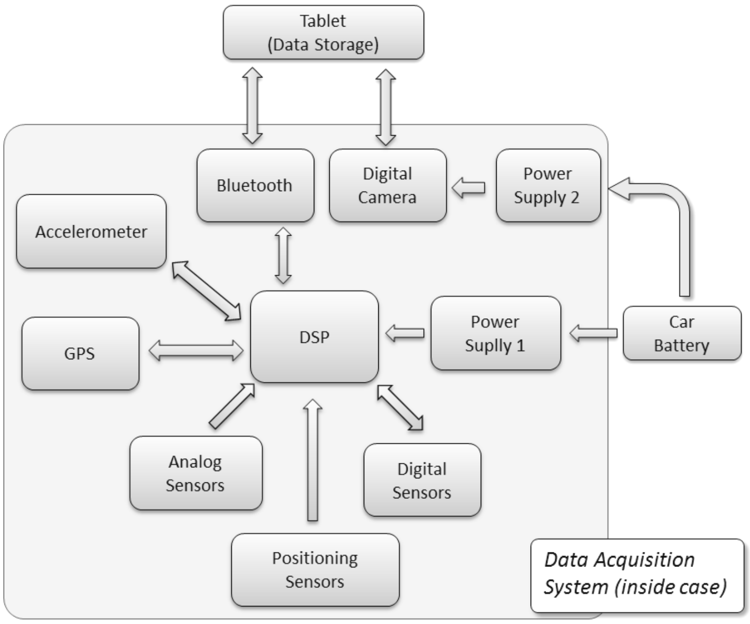


Figure 2: Data acquisition equipment architecture (extracted from [2]).

According to Keeffe [17], the radiation emitted by a specific type of lamp is composed of different spectrum lines, creating a spectrum signature, which depends on the lamp's chemical composition. Therefore, it's valuable information to collect the radiation emitted by the lamps, and, thus, the system was built with two types of radiometric sensors with different spectral responses, analog and digital sensors, totaling 13 sensors. The analog sensors are simple off-the-shelf photodiodes and phototransistors that transduce the lighting information in electrical information (voltage or current). On the other hand, the digital sensors are integrated circuits that provide the lighting information in a digital form, i.e., these devices have an embedded analog-to-digital converter. Thus, the choice of each device was based on the emission characteristics of the lamps improving the detection capacity of the proposed system. Table 1 summarizes the main photometric characteristics of the chosen sensors where each sensor provides a functional feature to the database.

Sensor ID	Wavelength of Maximum Sensitivity (nm)	Spectral Range of Sensitivity (nm)
1	560	460 ... 660
2	810	460 ... 1100
3	940	800 ... 1200
4	570	350 ... 1100
5	570	400 ... 900
6	880	730 ... 1100
7	640	300 ... 1150
8	650	300 ... 1100
9	800	500 ... 1100
10	480	380 ... 570
11	540	420 ... 630
12	620	580 ... 680
13	670	380 ... 680

Table 1: Photometrical features of the radiometric sensors

Another crucial piece of information that greatly influences the sensor's response is the magnitude of the lamp's spectrum lines which is directly correlated to the power of the lamps. Despite this correlation, the magnitude is extremely sensitive to external conditions such as ambient temperature, pole height, and boom angle. For this reason, predicting the power of a lamp is a more complex task and, consequently, requires more complex strategies.

Given this, the theoretical recognition of each parameter's influence on the magnitude of the spectral lines is an arduous task since all variables are correlated. Therefore, in order to provide more useful data to the model, a camera was coupled to the system in order to capture images of the lamps.

The model of the camera used is a Canon EOS Rebel T2i which consists of a professional camera with an impressive zoom ratio, a high resolution image, an autofocus system and the DIGIC 4 image processor, which helps process images quickly and efficiently. Given these features, the model was chosen to acquire the image data for the system.

In addition to the sensors and camera, the data acquisition equipment incorporates positioning sensors, which are light sensors that indicate to the operator whether the vehicle is situated directly beneath the luminaire. An accelerometer, which calculates the ground slope, and a Global Positioning System (GPS), which provides the geographical location of the pole, are also present.

Except for the digital camera, all the subsystems are managed by a digital signal processor (DSP), which collects their information and sends it by a bluetooth device to an external data storage system allocated inside the vehicle, such as a tablet computer. The DSP has an analog-to-digital converter that digitalizes the information provided by the analog sensors. The electronic device built can be seen at Figure 3.

(a) (b)

Figure 3: A picture of the implemented electronic board: (a) Top face: light sensors; (b) Bottom layer: main modules and analog conditioning circuits (extracted from [4]).

Then, the software processes the data gathered from the sensors, the camera, and the GPS, and identifies the lamp's location, type, and wattage, sending the new information to the company database.

4. THEORETICAL BACKGROUND

The upcoming section serves as the theoretical foundation addressing the details and features of the data gathered by the system, the classifiers used for intelligent processing, and the importance of feature selection in enhancing the system's efficacy. These components collectively underpin the system's efficiency and accuracy, contributing to a comprehensive understanding of its functioning.

4.1 Image Processing

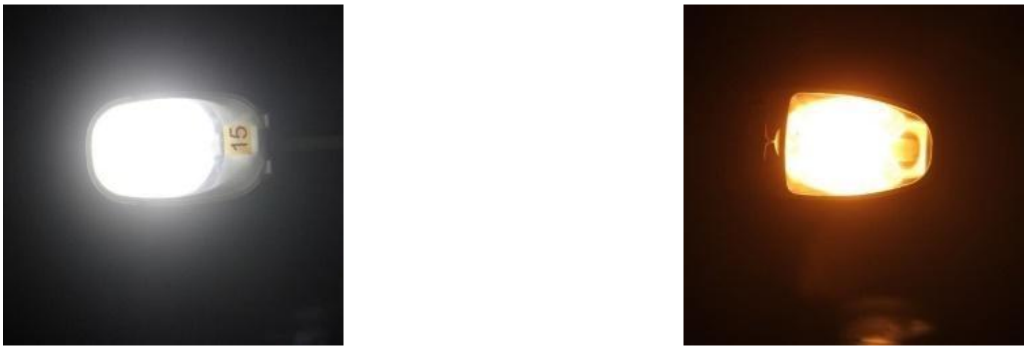
Image processing techniques were used to extract valuable features from the images collected by the system's digital camera. But, from the rough images, it's already possible to obtain some information. Figure 4 shows instances of the three lamp types present in the database. It can be noticed that each type has different colors, shapes, and forms. Moreover, in Figure 5, the lack of visual similarity between different types of lamps with the same wattage reaffirms the complexity of detecting patterns to distinguish the lamp's power. Therefore, visual descriptors were used to extract information about the texture and shape as separators for the problem.

(a) (b) (c)

Figure 4: Photos of three different types of street light bulbs: MV (a), MH (b), and HPS (c)

Before extracting the descriptors, another image processing technique was used, the threshold segmentation method [18], with the objective of separating the lamps from the image background. This process provided two more representations of the images, with gray and binary scale, shown in Figure 6 .

Therefore, using the three different descriptors, it was possible to summarize the image information in a set of 22 attributes: 10 Fourier descriptors (obtained from the binary representation), 7 Hu descriptors (obtained from the grayscale representation) and, 5 Haralick descriptors (also obtained from the grayscale representation).



(a) (b)

Figure 5: Photos of two different 150W street light bulbs: MH (a) and HPS (b)

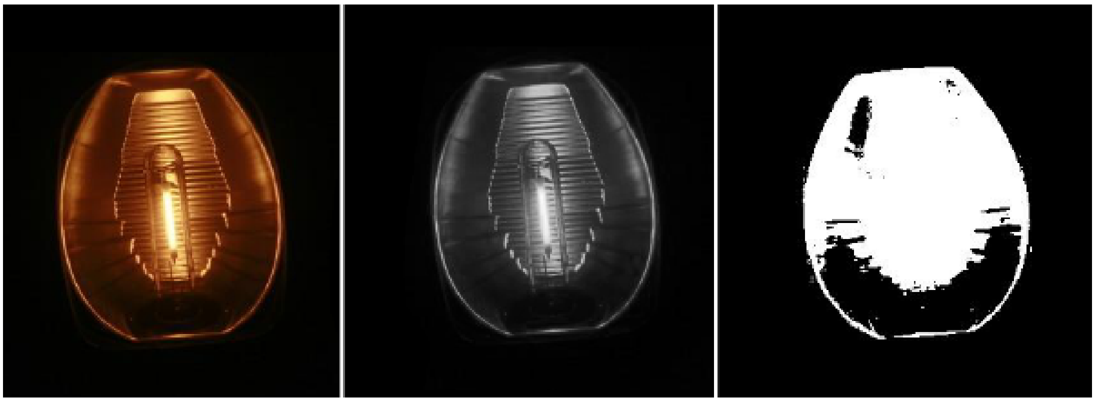


Figure 6: Original photo of a PLF unit (left), grayscale image (center) and its representation in binary scale (right)

The Fourier Transform is a tool for signal processing, in which each of its coefficients provides information about the studied object's outline and shape. And once the object's contour pixels can be represented as a signal, it's possible to obtain its contour points through a function given by a set of N samples in the frequency domain, denoted by $F(n), n = 0, 1, 2, \ldots, N-1$ and defined in Equation 1.

$$F_n = \frac{1}{N}\sum_{k=0}^{N-1} f(k) \cdot e^{-j2\pi nk/N} \tag{1}$$

Besides the large number of coefficients given by the Fourier Transformation, a subset is enough to provide valuable data about the contour and shape of the object. Thus, only the first 10 normalized Fourier coefficients, from [14], were used to create features.

On the other hand, Haralick descriptors [16] uses a statistical approach to extract texture information from an object, using the relationship between the digital image and its grayscale distribution. Thus, the analysis of the special distribution of pixels and their respective gray levels is made through the Gray Level Co-occurrence Matrix (GLCM), which is given by Equation 2:

$$C_{x,y}^{i,j} = \sum_{p=1}^{n}\sum_{q=1}^{m} \begin{cases} 1, & \text{if } |i-p+x| \leq p \text{ and } |j-q+y| \leq q \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

From the Haralick descriptors proposed in the original model [19], only the most promising for distinguishing the different classes of lamps power were used: energy, entropy, contrast, correlation, and homogeneity.

Another approach to feature extraction from images is the statistical moments (also called moments of an image). These moments, given by Equation 3, can provide information about the spatial distribution of points in a bi-dimensional image. The set of seven invariant moments used in this work to describe an image is known as the Hu descriptors, as defined in [15].

$$M_{ij} = \sum_{x,y} x \cdot y \cdot x_i \cdot y_j \cdot I(x,y) \tag{3}$$

4.2 Classification Algorithms

A classifier can be defined as a mathematical function that maps input data to a given category. Accordingly to [20], classification algorithms implement classifiers by analyzing a set of input data instances whose category membership is known. As previously described, each type of lamp and power is treated as a class, having features provided by the light sensors and image descriptors. Therefore, classifiers can be used in this scenario in order to identify the lamp class. This study uses both traditional and cutting-edge methods, such as Deep Learning, which are depicted in this subsection.

4.2.1 Traditional Machine Learning

K-Nearest Neighbour (KNN) is a non-parametric object classification method [19], based on the closest training examples in the feature space. Thus, an instance is ranked by the majority vote of its neighbors, with the object being assigned the class most common among a positive integer k of nearest neighbors.

There are several Decisions Trees (DT) inducing algorithms. This work uses the Python implementation of the *C4.5* algorithm, developed by [12]. *C4.5* finds high precision hypotheses through a specific and characteristic method, looking for pruning rules from the decision tree built during the training stage.

Created by Breiman [21], Random Forest (RF) is an algorithm that uses an ensemble of individual decision trees, each one built according to a random parameter. The great advantage of this technique is that the forest of trees protect itself from the individual errors of one. Random Forest (RF) combines multiple decision trees to make more accurate predictions. It works by creating a multitude of decision trees during the training phase and then aggregating their outputs to arrive at a final prediction. Each tree is trained on a random subset of the data and features, reducing the risk of overfitting and increasing the model's robustness. Random Forests are known for their versatility, high predictive accuracy, and ability to handle both classification and regression tasks.

Introduced by Friedman [22], Extreme Gradient Boosting (XGB) uses a technique that allows for a combination of gradient enhancement and RF. This algorithm works similarly to RF, with the difference that its trees are built sequentially, not parallel. In this way, each added tree minimizes the errors made by the previous one.

Suport Vector Machine (SVM) presented in [23] is an algorithm that finds a hyperplane that best separates features into different domains. This method, like KNN, is based on instance metrics used for predicting the classification.

The Multi Layer Perceptron (MLP) described in [24] is a feedforward artificial neural network. Typically, a neural network consists of processing units comprising an input layer, one or more hidden layers, and an output layer. The input signal propagates through the network in a direct layer-by-layer basis.

4.2.2 Deep Learning

The Convolutional Neural Network (CNN) is a variation of the MLP algorithm. According to Le Cunn [25], like traditional algorithms for processing visual data, CNN can apply filters to images, preserving the neighborhood association between its pixels throughout the processing of its entire network. Thus, this type of network is used in several problems of pattern recognition and detection in images and videos.

This network is composed of three types of layers, the convolutional, pooling, and fully connected layers, and their traditional architecture can be seen at Figure 7. The convolutional layer, seen at Figure 8, involves sliding a small filter (also known as a kernel) over the input data to extract local features by performing element-wise multiplications and aggregating the results to create feature maps. The pooling layer reduces the spatial dimensions of feature maps by selectively down-sampling, typically through operations like max-pooling or average-pooling, to retain essential information while reducing computational complexity. At the end of the model, the fully connected layer connects all neurons from the previous layer to every neuron in its layer, allowing for complex feature combinations and final decision-making in a neural network.

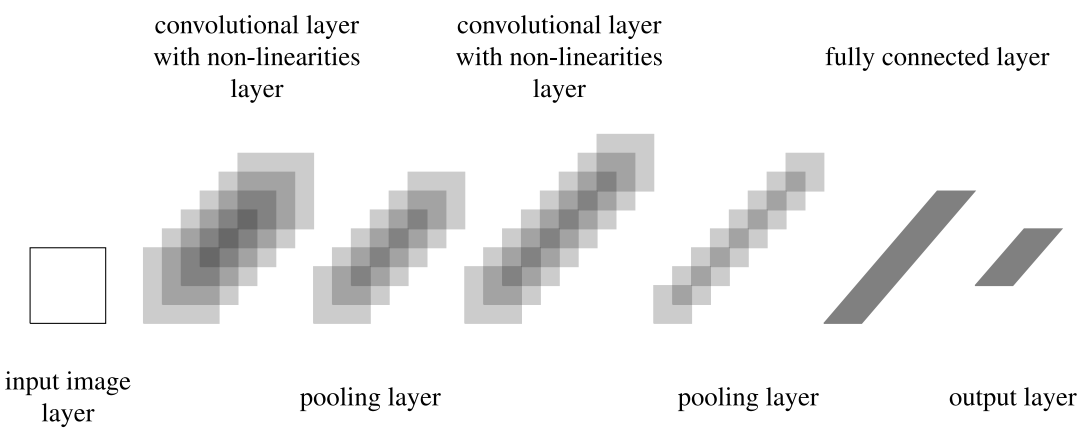


Figure 7: The architecture of the original convolutional neural network, as introduced by [26], alternates between convolutional layers including hyperbolic tangent non-linearities and subsampling layers. In this illustration, the convolutional layers already include non-linearities and, thus, a convolutional layer actually represents two layers. The feature maps of the final subsampling layer are then fed into the actual classifier consisting of an arbitrary number of fully connected layers. The output layer usually uses softmax activation functions.

The model is fed with the image as input, and the probabilities of each class are returned as output, where the class referring to the highest probability value will be chosen as the model prediction for the evaluated image. Thus, unlike the traditional machine learning approaches, CNN will exclusively use the original images collected by the system proposed by [2–4].

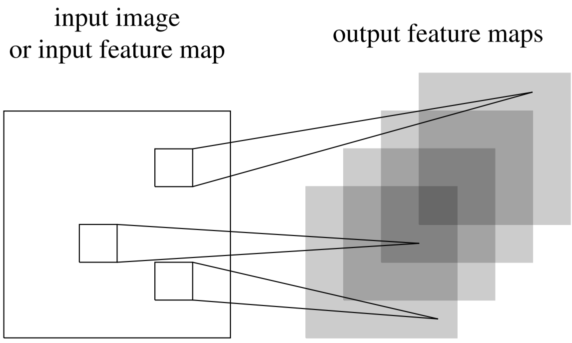


Figure 8: Illustration of a single convolutional layer. If it is the first convolutional layer, its input is a image. If not, its input is a feature map generated by the previous convolutional layer after applying different filters to yield the output feature map.

5. EXPERIMENTAL METHODOLOGY

This section describes the public lighting dataset, the resampling methodology, and the chosen metric for evaluating the performance of the machine learning models. This information is essential for understanding the reliability of our results.

5.1 Public Lighting Dataset

The dataset contains three types of lamp technologies: mercury vapor (MV), high-pressure sodium (HPS), and metal-halide (MH). These types of lamps were chosen due to the predominance of these technologies in the selected cities for the collection. As previously described, each type of lamp and wattage is treated as one class. Each dataset sample has features provided by the radiometric sensors and image descriptors.

Experiments were performed using data collected exclusively gathered in the field, as this particular data poses a more intricate classification challenge and better reflects the actual conditions encountered in urban streets and thoroughfares.

The dataset is composed by 297 samples distributed in 9 different classes, as shown in Table 2. The class column presents the classes used in this work, the type column indicates the type of the lamp (HPS, MH or MV), the Wattage column indicates the lamp wattage, the support column shows the number of samples of each class and the percentage column shows the proportion of samples of each class. As one can see, the dataset distribution is sufficiently balanced.

Table 2: Class instances distribution summary

Class	Type	Wattage (W)	Support	Percentage (%)
HPS070	HPS	70	30	10.1%
HPS100	HPS	100	32	10.8%
HPS150	HPS	150	35	11.8%
HPS250	HPS	250	33	11.1%
HPS400	HPS	400	37	12.5%
MH150	MH	150	23	7.7%
MH250	MH	250	49	16.5%
MH400	MH	400	37	12.5%
MV125	MV	125	21	7.1%
Total			297	100.0%

All features of the dataset were normalized by applying a minmax approach as indicated by Equation 4.

$$z_i = \frac{x_i - \min(x)}{\max(x) - \min(x)} \tag{4}$$

5.2 Resampling Method

In machine learning, models are trained on labeled data to make accurate predictions. During training, models adjust their parameters to minimize errors on the training dataset. Afterward, models are tested on another independent dataset to provide an unbiased assessment of their real-world performance, ensuring their ability to make accurate predictions in practical applications. Resampling methods must prioritize ensuring the independence of both the training and test sets. Any violation of this crucial aspect may lead to a test leak, introducing bias into the model and significantly compromising its performance when applied to new, real-world data [27].

While ensuring independence between the sets is vital, it is important to note that it does not guarantee improved model generalization. This is because the distribution of the data division can still introduce bias into the learning process, if only one split is performed. In order to get a more robust resampling method to evaluate model performance, it is common to use the k-fold cross-validation technique. This process partitions the data into k folds, subsets, iteratively training the model k times. In each iteration, one fold serves as the testing set, while the remaining $k-1$ folds are used for training. This process provides a more comprehensive assessment of the model's performance, reducing the risk of overfitting and yielding a reliable estimate of its generalization ability across k different test data subsets. To replicate real-world scenarios and preserve the distribution of the original dataset within these folds, the data division is generally carried out using a stratified approach.

However, k-fold cross-validation technique still is subjected to biased results. This may happen because of the hyperparameter tuning, which is an essential step in machine learning as it fine tunes model settings to optimize performance. Tuning these settings ensures a model's adaptability to diverse datasets, customizes it for specific tasks, and enables comparative analysis of different models. By striking the right balance between hyperparameters, machine learning models can achieve higher accuracy, better generalization, and cost-effective resource utilization, making this process an essential component of model development. Nonetheless, whenever using the k-fold cross-validation technique, typically the chosen hyperparameters values are those that maximize the average performance on the k test sets. So, the testing dataset influences the training of the models, introducing bias and generating optimistic results that will not be repeated to new, real-world data.

In order to avoid biased results, the nested cross validation resampling method is used in this work. It involves two levels of cross-validation: an outer and an inner loop. In this method, the data is divided in training, validation and test sets. The validation dataset is used for hyperparameter tuning. The outer loop performs a conventional k-fold cross-validation for estimating the model performance. However, for each test set of the outer loop, it is also performed another k-fold cross validation in the inner loop, where one fold is selected for validation in each iteration. In the inner loop, a grid search technique [28] is used to select the best hyperparameters for the model on the training set without being influenced by the test set as shown in Figure 9.

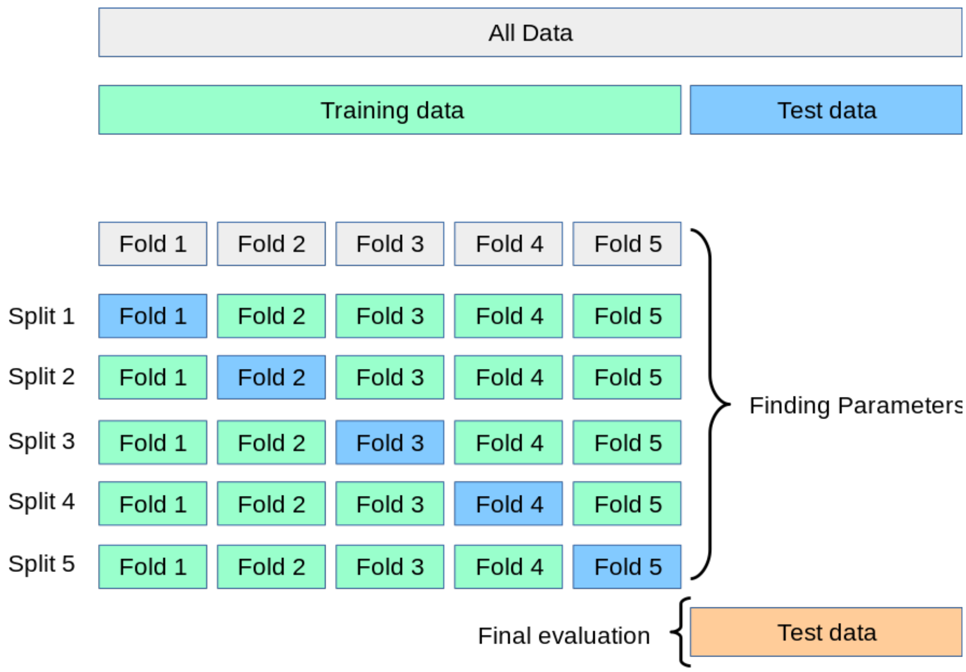


Figure 9: A 5-fold cross validation inner loop (extracted from [29]).

In the grid search, you define a range of possible values for each hyperparameter of interest. The technique then exhaustively tests all possible combinations of these hyperparameters using k-fold cross-validation to validate each combination's performance. This involves training and validating the model with each unique set of hyperparameters to determine which combination produces the best results in terms of a chosen evaluation metric. The model's performance, assessed using an evaluation metric, solely informs the selection of the best hyperparameters to be applied in the test set during the outer cycle. The hyperparameters values of each model used in the experiments are shown in tables 3 to 8.

Table 3: Range of hyperparameter values for the KNN algorithm

Hyperparameters	**Range of values**
number of neighbors	1, 3, 5, 7, 9
weights	uniform, distance
algorithm	auto, ball_tree, kd_tree, brute

Table 4: Range of hyperparameter values for the DT algorithm

Hyperparameters	Range of values
max_depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 15, 20, 25, 30
max_features	10, 11, 13, 15, 17, 19, 20, 21, 23, 25
min_samples_split	32, 64, 128, 256
min_samples_leaf	32, 64, 128, 256

Table 5: Range of hyperparameter values for the RF algorithm

Hyperparameters	Range of values
max_depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 15, 20, 25, 30
max_features	10, 15, 20, 25
n_estimators	50, 100, 300, 500, 1000, 2000, 3000

Table 6: Range of hyperparameter values for the XGB algorithm

Hyperparameters	Range of values
max_depth	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 13, 15, 20, 25, 30
learning_rate	0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95
n_estimators	5000

Table 7: Range of hyperparameter values for the SVM algorithm

Hyperparameters	Range of values
C	0.1, 1, 10, 30, 50, 70, 90, 100, 300, 500, 700, 1000, 2000
gamma	1, 0.1, 0.01, .001
kernel	rbf, poly, sigmoid

Table 8: Range of hyperparameter values for the MLP algorithm

Hyperparameters	Range of values
activation	identity, logistic, tanh, relu
solver	sgd, adam
learning_rate	constant, invscaling, adaptive
max_iter	5000
tol	0.001
learning_rate_init	0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65, 0.7, 0.75, 0.8, 0.85, 0.9, 0.95
momentum	0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9
hidden_layer_sizes	(9,), (14,), (15,), (16,), (17,), (18,), (19,), (20,), (21,), (24,), (25,), (27,), (28,), (30,)

5.3 Evaluation

The accuracy metric is the most frequent performance measure used in classification tasks to evaluate the overall correctness of predictions made by a machine learning model. It is defined as the ratio of correctly predicted instances to the total number of instances in the dataset. The mathematical expression for accuracy is given by Equation 5.

$$\text{Accuracy} = \frac{\text{Number of Correct Predictions}}{\text{Total Number of Predictions}} \tag{5}$$

It is expressed as a value in the range $\{0, 1\}$, where 1 indicates that all predictions were correctly classified, and 0 indicates complete misclassification. It is a straightforward and intuitive metric that is widely used in the literature for balanced datasets. Hence, this work uses this metric to evaluate the performance of the models.

6. RESULTS

This section first describes the results obtained by the conventional machine learning methods and later describes the results of the deep learning approach.

In an effort to comprehensively assess the information sources and determine if they can complement one another, the experiments were organized into three distinct data domains: data derived from sensor information, data originated from image descriptors, and a combination of sensor data and image descriptors data. This approach aimed to evaluate the knowledge extracted from each source effectively and how they complement each other. The results of the traditional machine learning classifiers are shown in Table 9.

Table 9: Average accuracy and Confidence Interval (95%) of machine learning classifiers

Algorithm	Sensors	Images	Sensors & Images
KNN	$0.51\ (0.48-0.55)$	$0.27\ (0.24-0.30)$	$0.42\ (0.39-0.46)$
DT	$0.50\ (0.47-0.53)$	$0.39\ (0.36-0.42)$	$0.52\ (0.48-0.55)$
SVM	**0.86 (0.83 - 0.88)**	$0.61\ (0.59-0.63)$	**0.83 (0.81 - 0.85)**
RF	$0.79\ (0.76-0.82)$	**0.65 (0.61 - 0.69)**	**0.83 (0.80 - 0.86)**
XGB	$0.76\ (0.73-0.78)$	$0.62\ (0.59-0.64)$	$0.82\ (0.79-0.85)$
MLP	$0.67\ (0.63-0.70)$	$0.51\ (0.47-0.54)$	$0.79\ (0.76-0.82)$

One may see that the more robust methods (SVM, RF, XGB) indeed delivered the best performance in the three different scenarios. Best results are indicated in bold text in Table 9. The SVM method achieved the highest average accuracy with sensors only data, while the RF method achieved the highest average accuracy with images descriptors only data. SVM and RF achieved the highest average accuracy with both sensors and image descriptors data. The experiment results show that sensors data is more informative than image descriptors data by far. However, most machine learning models slightly improved their performance with the two sources of data. Unexpectedly, the overall best performance was gathered by the SVM method using the sensors data only.

Although hyperparameter tuning has many advantages, it may be time-consuming and computational expensive. Hence, searching for the best set of parameters of long-training models tends to be impracticable. Therefore hyperparameter tuning was not performed in the deep learning approach. A single CNN comprising five convolutional layers, each succeeded by a max-pooling layer and one last fully connected layer were used. Each max-pooling layer reduces the input layer by half. Despite the lack of hyperparameter tuning, these models are robust, capable of directly processing raw images, shifting the responsibility of feature extraction from the developer to the network. This allows the network to learn hierarchical representations of the data, resulting in less biased analyses. The proposed CNN obtained average accuracy of 0.62 with a Confidence Interval (95%) of 0.43–0.80 processing the raw images only. This result is similar to those obtained by the more robust machine learning models using the image descriptors data only, but is significantly lower to those obtained using sensors data only and sensors and image descriptors data.

7. CONCLUSIONS

In this paper, a methodology that enhances the intelligent model proposed by [2–4] for the identification of street lighting lamp type and wattage is presented. Two approaches were addressed. The first approach involved the application of traditional machine learning classifiers (KNN, DT, RF, SVM, XGB, and MLP) to three distinct domains: radiometric sensor data, image descriptors data, and a combination of both. The second approach utilized a deep learning convolutional network (CNN) that directly processed the raw images captured from the lamps. In addition, the validation process were enhanced by implementing a more robust technique, known as nested cross-validation. This approach yields more reliable results when compared to those presented in previous studies.

As expected, the more robust methods (SVM, RF, XGB) indeed achieved the best performance in the three different scenarios. Results also showed that radiometric sensor data is more informative than image descriptors data for the classification task, although there is indication that they may be complementary. The deep learning approach achieved results similar to the more robust traditional methods when using the image descriptors only. The deep learning approach was not competitive with the traditional most robust methods when using the sensor data only or using the both data sources.

Experimental results confirmed that the proposed method is quite effective for reducing the commercial losses caused by non updated public lighting databases. The experiment best result was achieved by the SVM method using the radiometric sensors data only. It achieves an average accuracy of 0.86. Since the public lighting dataset has 9 relatively balanced classes, the achieved accuracy is significantly high. Human inspectors do not have similar performance just by looking under the street lamp.

Possible topics in future works include investigating more robust and sophisticated architectures of deep learning CNNs including influential models like ResNet (Residual Networks) [30] and AlexNet [31], and the use of pre-trained CNNs fine tuned by the public lighting dataset. Another possible future work consists of expanding the public lighting data set by collecting new data instances and also obtaining data from new features that may be relevant information for the classification, such as the pole height.

REFERENCES

[1] "Manual para elaboração do Programa Anual de Combate ao Desperdício de Energia Elétrica da ANEEL", 1999. Last accessed 23 August 2023.

[2] G. M. Soares, A. G. B. Almeida, R. M. Mendes, E. C. Teixeira, H. A. C. Braga, M. N. Machado, R. S. Broetto, M. V. H. B. Castro, H. O. G. Filho, F. M. Varejão, J. G. P. Filho, A. B. Candeia and R. A. A. Sousa. "Detection of street lighting bulbs information to minimize commercial losses". In *2013 Seventh International Conference on Sensing Technology (ICST)*, pp. 895–900, 2013.

[3] G. M. Soares, A. G. B. Almeida, R. M. Mendes, E. C. Teixeira, H. A. C. Braga, M. N. Machado, R. S. Broetto, M. V. H. B. Castro, H. O. Gomes Filho, F. M. Varejão, J. G. Pereira Filho, A. B. Candeia and R. A. A. Sousa. "Performance evaluation of a sensor-based system devised to minimize commercial losses in street lighting networks". In *2014 IEEE International Instrumentation and Measurement Technology Conference (I2MTC) Proceedings*, pp. 336–341, 2014.

[4] G. M. Soares, A. G. B. Almeida, R. M. Mendes, E. C. Teixeira, H. A. C. Braga, M. N. Machado, R. S. Broetto, M. V. H. B. Castro, H. O. G. Filho, F. M. Varejão, J. G. P. Filho, A. B. Candeia and R. A. A. Sousa. *On the Use of Light Sensors and Pattern Recognition Techniques for Automated Detection of Street Lighting Lamps*, pp. 67–92. Springer International Publishing, Cham, 2015.

[5] M. Dwiyaniti, A. B. Kusumaningtyas, S. Wardono, K. Sri Lestari and Tohazen. "A Real-time Performance Monitoring of IoT based on Integrated Smart Streetlight". In *2022 6th International Conference on Electrical, Telecommunication and Computer Engineering (ELTICOM)*, pp. 131–135, 2022.

[6] M. Kanthi and R. Dilli. "Smart streetlight system using mobile applications: secured fault detection and diagnosis with optimal powers". *Wireless Networks*, vol. 29, no. 5, pp. 2015–2028, July 2023.

[7] M. Liu, B. Zhang, T. Luo, Y. Liu, B. A. Portnov, T. Trop, W. Jiao, H. Liu, Y. Li and Q. Liu. "Evaluating Street Lighting Quality in Residential Areas by Combining Remote Sensing Tools and a Survey on Pedestrians' Perceptions of Safety and Visual Comfort". *Remote Sensing*, vol. 14, no. 4, pp. 826, January 2022. Number: 4 Publisher: Multidisciplinary Digital Publishing Institute.

[8] I. Mavromatis, A. Stanoev, P. Carnelli, Y. Jin, M. Sooriyabandara and A. Khan. "A dataset of images of public streetlights with operational monitoring using computer vision techniques". *Data in Brief*, vol. 45, pp. 108658, 2022.

[9] R. S. Broetto, F. M. Varejão, J. Gonçalves Pereira Filho and A. Bernadi Candeia. "Automated image recognition of public lighting luminaries". In *2013 IEEE 20th International Conference on Electronics, Circuits, and Systems (ICECS)*, pp. 657–660, 2013.

[10] R. S. Broetto and F. M. Varejão. "Heterogeneous feature models and feature selection applied to detection of street lighting lamps types and wattages". In *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, pp. 933–938, 2016.

[11] D. K. D.W. Aha and M. Albert. "Instance-based learning algorithms". *Machine Learning*, vol. 6, no. 1, pp. 37–66, 1991.

[12] J. Ross Quinlan. *C4.5: Programs for Machine Learning*, volume vol 16. Morgan Kaufmann Publishers, 1994.

[13] D. Rumelhart and J. McClelland. *Learning Internal Representations by Error Propagation*. MIT Press, Cambridge, MA, 1987.

[14] T. R. Crimmins. "A Complete Set of Fourier Descriptors for Two-Dimensional Shapes". *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 12, no. 6, pp. 848–855, 1982.

[15] M.-K. Hu. "Visual pattern recognition by moment invariants". *IRE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187, 1962.

[16] R. M. Haralick, K. Shanmugam and I. Dinstein. "Textural Features for Image Classification". *IEEE Transactions on Systems, Man, and Cybernetics*, vol. SMC-3, no. 6, pp. 610–621, 1973.

[17] W. Keeffe. "RECENT PROGRESS IN METAL HALIDE DISCHARGE-LAMP RESEARCH." *IEE Proceedings A: Physical Science. Measurement and Instrumentation. Management and Education. Reviews*, vol. 127, no. 3, pp. 181 – 189, 1980.

[18] T. Ridler and S. Calvard. "PICTURE THRESHOLDING USING AN ITERATIVE SLECTION METHOD." *IEEE Transactions on Systems, Man and Cybernetics*, vol. SMC-8, no. 8, pp. 630 – 632, 1978. Cited by: 1867.

[19] E. Fix and J. L. Hodges. "Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties". *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, pp. 238–247, 1989.

[20] S. Kotsiantis. "Supervised machine learning: A review of classification techniques". *Informatica (Ljubljana)*, vol. 31, no. 3, pp. 249 – 268, 2007.

[21] L. Breiman. "Random Forests". *Machine Learning*, vol. 45, no. 1, pp. 5–32, October 2001.

[22] J. H. Friedman. "Greedy function approximation: A gradient boosting machine." *The Annals of Statistics*, vol. 29, no. 5, pp. 1189–1232, October 2001. Publisher: Institute of Mathematical Statistics.

[23] C. Cortes and V. Vapnik. "Support-Vector Networks". *Machine Learning*, vol. 20, no. 3, pp. 273–297, September 1995.

[24] D. E. Rumelhart, G. E. Hinton and R. J. Williams. "Learning representations by back-propagating errors". *Nature*, vol. 323, no. 6088, pp. 533–536, October 1986.

[25] Y. LeCun, Y. Bengio and G. Hinton. "Deep learning". *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015. Number: 7553 Publisher: Nature Publishing Group.

[26] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard and L. D. Jackel. "Backpropagation Applied to Handwritten Zip Code Recognition". *Neural Computation*, vol. 1, no. 4, pp. 541–551, December 1989. Conference Name: Neural Computation.

[27] T. Cover and P. Hart. "Nearest neighbor pattern classification". *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

[28] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and Édouard Duchesnay. "Scikit-learn: Machine Learning in Python". *Journal of Machine Learning Research*, vol. 12, no. 85, pp. 2825–2830, 2011.

[29] Sklearn. "Cross-validation: evaluating estimator performance". Available at `https://scikit-learn.org/stable/modules/cross_validation.html` (2023/09/20). Last accessed 20 September 2023.

[30] K. He, X. Zhang, S. Ren and J. Sun. "Deep Residual Learning for Image Recognition". In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778, June 2016. ISSN: 1063-6919.

[31] A. Krizhevsky, I. Sutskever and G. E. Hinton. "ImageNet classification with deep convolutional neural networks". *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, May 2017.