Anomaly detection using Computer Vision for fusion power generation

Bachelor's Thesis

Abraham Cano Aguilera

abraham.cano@estudiant.upc.edu

Image Processing Group
Signal Processing and Communications Group
UPC

Supervisors:
Prof. Dr. Josep Ramon Morros
Prof. Dr. Josep Ramon Casas

June 29, 2020

Acknowledgements

I would like to thank my thesis advisor Dr.Josep Ramon Morros of the Universitat Politècnica de Catalunya. The door to Prof. Morros office was always open whenever I ran into a trouble spot or had a question about my research or writing. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever he thought I needed it.

I would also like to appreciate master student Marti Cobos for his previous research,helpful advises and work provided. Without he, it would have been unimaginable the progress achieved. I would also like to acknowledge the technical staff,special mention to Josep Pujal for their constant assistance whenever it has been necessary.

Finally, I must express my very profound gratitude to my parents and to my personal circle for providing me with unfailing support and continuous encouragement throughout ny yeas of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Abstract

Through the years, nuclear fission, the most extended way to generate nuclear energy, has been subject of debate because its detrimental effects in the environment as well as our health. By now nuclear fusion is emerging as a most secure method to produce a clean power,and as soon the good results will achieve, it probably will be the energy that will feed our homes and countries.

The problem with this process is that physicists have several problems to recreate the necessary conditions to achieve a stable reaction, and regrettably this troubles in the procedure can cause serious damage in valuable devices.

Now computer vision becomes as a potential solution to address this issue. With the implementation of visible cameras as well as infrared cameras, it is possible generate a huge useful video sequences to analyze abnormal behaviour in the W7-X stellarator.

In this thesis will be studied the procedure to generate a complete thermal anomaly detector and classifier standing by spectrum visible cameras. It will be developed a system to extract a descriptive feature descriptors that generalize each detection, as well as a system to classify this features as anomaly or not. Finally it will be detailed the results of survey in order to reveal the most useful tools to deal with this type of images and therefore contribute to next related investigations.

Resumen

A través de los años, la fision nuclear, la forma más extendida de generar energia nuclear, ha sido objeto de debate a causa de sus perjudiciales efectos sobre el medio ambiente así como nuestra salud. En la actualidad, la fusión nuclear ha emergido como un metodo más fiable para producir una fuente de energia limpia que tan pronto como obtenga resultados, alimentara probablemente nuestros paises y hogares.

El problema con este proceso radica en que los físicos involucrados tienen considerables problemas a la hora de recrear las condiciones necesarias para conseguir una reacción estable, y desgraciadamente estos impedimentos pueden causar daños severos en el reactor.

Ahora la visión por computador se ha convertido en una esperanzadora solución para abordar el problema. Con la implementación de cámaras infrarojas así como de espectro visible dentro del reactor, es posible generar extensas secuencias de video, que luego seran procesadas con el objetivo de detectar comportamiento inusual.

En esta tesis será estudiado el procedimiento necesario para generar un detector y clasificador de zonas calientes en el reactor apoyándonos en imagenes del espectro visible. Será desarrolado un sistema para extraer las características descriptivas de la imagen que mejor generalizen las secuencias, así como un sistema para clasificar estas caracteristicas como anomalias.Finalmente seran detallados los resultados del estudios, con el objetivo de poder mostrar las técnicas más utiles para procesar este tipo de imagenes y por tanto contribuir a futuras investigaciones.

Resum

A través dels anys, la fisió nuclear, la forma més estesa de generar energia nuclear, ha estat objecte de debat a causa dels seus perjudicials efectes sobre el medi ambient, així com la nostra salut. En l'actualitat, la fusió nuclear ha emergit com a mètode més fiable per produir una font d'energia neta, que tan aviat com obtingui resultats, alimentarà probablement els nostres països i llars.

El problema amb aquest procés rau en que els físics involucrats tenen considerables problemes a l'hora de recrear les condicions necessàries per aconseguir una reacció estable, i malauradament, aquests impediments poden causar danys severs en el reactor. Ara la visió per computador s'ha convertit en una esperançadora solució per abordar el problema. Amb la implementació de EDICAM 's dins el reactor, és possible generar extenses seqüències de vídeo, que després seran processades amb l'objectiu de detectar comportament inusual.

En aquesta tesis serà estudiat el procediment necesari per generar un detector i clasificador de regions sobrescalfades en el reactor. Es desenvoluparà un sistema per extraure les caracteritiques descriptives de la imatge que millor generalitzin les seqüències, així com un sistema per clasificar aquestes caracteristiques com anomalies. Finalment seràn detallats els resultats del estudi, amb l'objectiu de poder mostrar quines son les tècniques més útils per procesar aquest tipus de imatges i per tant contribuïr a futures investigacions.

Contents

Acknowledgements **i**

Abstract **ii**

Resumen **iii**

Resum **iv**

1 Introduction **1**

1.1 Objectives of the thesis 2

1.2 Requirements and specifications 3

1.3 Organization 3

2 State of the art **5**

3 Feature extraction techniques **8**

3.1 Gradient Based Approach 9

3.1.1 Histogram of oriented gradients (HOG) 9

3.2 Local binary patterns (LBP) 10

3.3 Key Points Approach:Sift 11

3.3.1 SIFT 12

4 Feature selection techniques **15**

4.1 Filter Based methods 15

4.1.1 Pearson Correlation 15

4.1.2 Chi-Squared 15

4.2 Wrapper-based methods 15

4.2.1 Recursive Feature Elimination 15

4.3 Embedded methods 15

4.3.1 Lasso 15

4.3.2 Tree-Based . . . 15

5 Classifiers **16**

5.1 Machine learning classifiers . . . 16

5.1.1 Supported vector machines (SVM) . . . 16

5.1.1.1 Linear Support Vector Machines: Hard-Margin case . . . 17

5.1.2 Decision tree classifiers . . . 18

5.1.2.1 Bagged decision trees: Random Forest Classifier 20

5.1.2.2 Boosted trees: Adaboost classifier . . . 20

5.1.2.3 Xgboost classifier . . . 21

6 Resampling methods **22**

6.1 Over-Sampling: SMOTE . . . 22

6.2 Under-sampling: Near-Miss . . . 23

7 Experiments and results **24**

7.1 Train procedure and assessment metrics . . . 24

7.1.1 Cross-Validation . . . 26

7.1.2 Assessment metrics . . . 26

7.2 Image descriptors and feature extraction . . . 27

7.3 Gradient global approach(HOG) . . . 28

7.4 Local key point approach(LBP) . . . 28

7.5 Classifiers . . . 29

7.6 Imbalanced data techniques . . . 29

7.7 Final results . . . 29

8 Conclusions **32**

9 Future Work **33**

Bibliography **34**

List of Figures

1.1 Project Gantt diagram 4

2.1 W7-X Image system overview. Extracted by [1] 5
2.2 W7-X Image Scene Model. Extracted from [1] 6
2.3 Temperature evolution of PFC. Extracted from [2] 7
2.4 Surface temperature evolution. Extracted from [2] 7
2.5 $norm_{t,T}$ and threshold. Extracted from [2] 7
2.6 Projection of the heat flux mapped on the horizontal and vertical divertor of W7-X. Extracted from [2] 7

3.1 LBP operationing 11
3.2 The DOG method in each octave 13
3.3 Local minimum or maximum detection. The pixel X is compared with its 26 adjacent pixels through each belonging and adjacent scale 13
3.4 At left the gradients of analyzed image. At right the key point descriptor exposed in [3] 14

5.1 Maximum-margin hyperplane and margins for a SVM trained with two classes samples. The samples contained in the margin are called support vectors 17
5.2 Maximum-margin linear decision function $D(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$ 18
5.3 Maximum-margin linear decision function $D(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b$ 19

7.1 Normal behaviour 25
7.2 Hot spot 25
7.3 Hot spot image processing 25
7.4 K-Fold operation 26

List of Tables

7.1 Comparison of the color descriptors with 2 classes approach . . . 30
7.2 Classifiers execution time 30
7.3 Comparison of the SMOTE performance between RFC and HOG 31

Nomenclature

CV	Computer Vision
DoG	Difference of Gaussians
$EDICAM$	Event Detection Intelligent Camara
FN	False Negatives
FP	False Positives
$FPPW$	False Positive Per Window
HOG	Histogram of Oriented Gradients
$ILBP$	Improved Local Binary Pattern
IPP	Max Planck Institute for Plasma Physics
IR	Infrared
LBP	Local Binary Patterns
LTP	Local Ternary Patterns
ML	Machine Learning
PCA	Principal Component Analysis
PFC	Plasma-Facing Components
ROC	Receiver Operating Characteristics
$SIFT$	Scale Invariant Feature Transform
TN	True Negatives
TP	True Positives

Chapter 1

Introduction

With the recent technological revolution, computers are presents for all aspects in our lives. This allow us to generate an enormous quantity of information and data that obviously is pretty difficult to process manually. In this context appears artificial intelligence, that aims to build smart machines capable to do tasks that normally require human intelligence.

Turning to the topic of the thesis, it is essential mention Computer Vision, a machine learning subfield, oriented to process the images in order to recreate the human behaviour processing the image inputs. Computer Vision has been heavily further exploited in fields such as autonomous driving, pedestrian detection or robotics, but there are vast amount of areas in which it would be useful but currently it is not much research done. At this juncture, plasma events detection in nuclear fusion reactors with CV techniques has a long way to go in order to get knowledge as well as equipment protection.

Many scientists are doing research in the nuclear fusion in order to get a clean source of energy (against nuclear fission). With the aim of obtain knowledge about the factors that can disrupt the nuclear fusion process, the Max Planck Institute for Plasma Physics(IPP) installed cameras inside the reactor to detect and classify anomalies during the nuclear fusion generation.

One of the principal issues during the nuclear fusion process is the overheating of the regions nears to the vacuum vessel that contains the plasma , also called hot spots. This hot spots can damage the cooling capabilities of the reactor and in the extreme cases may even lead to the superheating of the system. Therefore it is imperative detect their on time with the aim of stop the nuclear reaction, therefore AI becomes as a essential tool to detect,track and classify plasma events equal or better than humans.

1.1 Objectives of the thesis

The main tasks of this thesis are the detection,tracking,feature extraction and classification of the anomaly behaviour inside the reactor. The first two points was investigated and implemented by the master student Marti Cobos in consideration of the biggest challenge of this project is the absence of labeled sequences, therefore it is necessary the conducting of a manual labelling, which is very time-consuming or conversely, the implementation of a semi-automatic bright spot. This bright spot detector implemented by Marti is based on a sequence of mathematical morphology operations [4], that are very useful to detect the bright local regions in a image, in along with the connected components algorithm [5] to detect those areas, and the non maximum suppression algorithm that aims to remove the overlapping detections [6]. The problem with this detector is that it classifies all detections as hot spots, whilst only a small portion of them really are (about 5%). Therefore, it is essential the development of a system capable to reduce the number of false positives (section 6.1), with the aim of achieving automatic hot spot detector.

Once introduced the principle goal of the thesis, it is clear that the objectives of the thesis will be adquire the necessary knowledge of techniques behind a object recognition system. Consquently, will be the following:

1. **Data integration:** As mentioned above, this project needs external data provided from IPP,as well as Marti's source code. In consequence, it is challenging do a correct data integration in order to learn and take advantage from previously used methods.

2. **Feature extraction:** All the object recognition systems has to generalize the objects detected by way of some characteristics,known as image descriptors. It would be essential get the theorical knowledge of this methods as well as how implement its with the python computer vision package.

3. **Machine learning classifiers:** Once the object features have been extracted, it is necessary label those samples as hot spot or not (binary case). Machine learning classifiers are the solution ad hoc, therefore it is needed which classifiers are more appropiate for this CV task just as their implementation with the help of python science libraries.

In addition to this technical objectives they will also be the acquirement of abilities related to the thesis writing and research such as understand scientific publications, do a correct bibliographic study or learn how express the survey results.

1.2 Requirements and specifications

The project is part of novel investigation about the thermal events in the W7-X, thereupon it has no sense establish a hard requirements to the final system. Despite this, it is possible set some that will help its to keep on track and to know how close we are to reaching the final goal. The final software must be able to:

- **R1:** Generalize as well as possible the sequences provided by the IPP. The model will have to adjust to the future sequences not only the current.
- **R2:** Be robust against noise in the images. Clearly, the images inside the reactor are noisy, therefore the system should be capable to calibrate the image noise.
- **R3:** Be robust against apparent motion (illumination could change in each sequence, camera movement...).
- **R4:** Be robust against overfitting.Numerous hot spots has nearly identical features within the same sequence,so it becomes vital get results with unseen data with the aim of be able to guarantee that the model developed is appropriate.

Related to specifications, the project is a innovative investigation, so it does not have clear specifications. Despite this, it can be defined a few specifications that allows its to show the performance of the model developed:

- **S1:** The baseline performance metrics, for instance random classifier, should be improved.
- **S2:** The classifier should have a better recall than precision in order to avoid FP. Therefore, it should be prioritized the detection of spots as hotspots (Huge TP rate) although there may be lots of non hot spots detected as hot posts, better known as FN.

1.3 Organization

As can be seen in the figure 1.3, it is possible breaking the project down into steps.

The first part of the project consists in to make a research and learning about the essential computer vision concepts needed for resolve the goal. Even though this task is more time-consuming in the beginning of the project, it continued present until the project end.

The next step was the data processing and integration of the system developed previously. This was the most critical part because in the case of lack of adequate results, it would not have been possible pursue with the rest of the project. Thus, it was established a middlestone at 9 April.

Following the Gantt diagram, the next stage was to learn how label the video sequences provided by IPP, just as the implementation of the feature extractor and classifier. This was the most technical part, so it required most of the project time.

Sequentially,once the classifiers were trained, the last part of the project revolved around the generation and assessment of the results. One time the results were obtained, they was reported in this thesis.

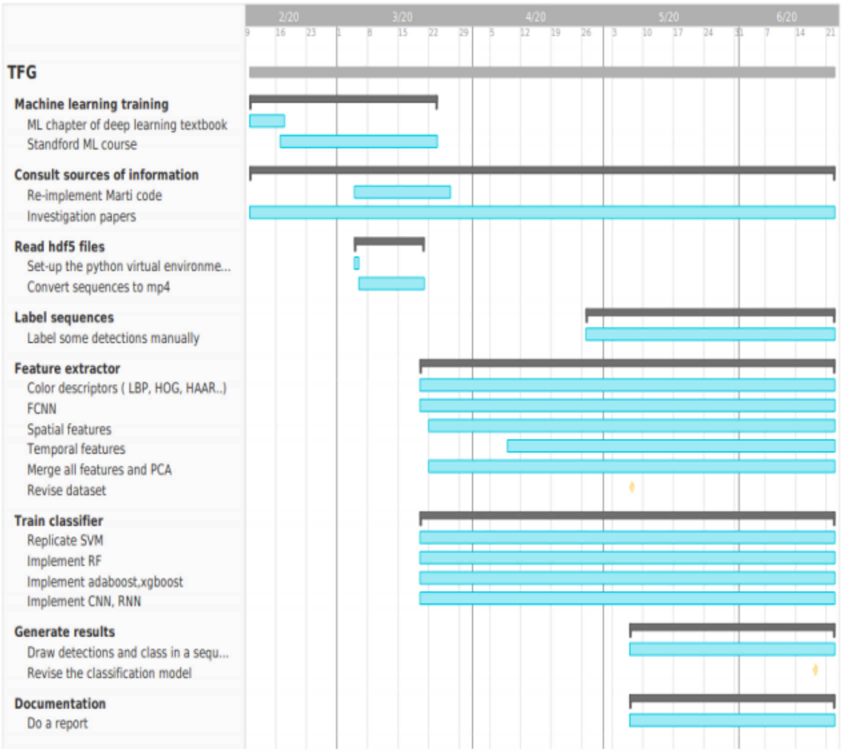


Figure 1.1: Project Gantt diagram

Chapter 2

State of the art

As explained in chapter 1, the detection of plasma events in the W7-X with CV techniques is a relative novel field of study since W7-X developing had finished at 2015. Nevertheless,there are a considerable number of surveys focused on plasma events detection by means of CV techniques. In this chapter, it will be addressed an overview of some of the most interesting research.

In [1], A.Puig et al. published and overview of the image diagnostic system for PFCs protection in the stellarator.
The image system of W7-X is composed of ten infrared cameras and ten video cameras. The video cameras,which are the ones that provide us the video sequences to analyze,also called EDICAMs offers a tangent vision of the plasmatic cavities. EDICAMs cover all the visible spectrum by means of CMOS LUPA-1300 sensor with a resolution of 1280x1024 pixels.
The software of the W7-X image aims to obtain and process the images from EDICAMs. The video sequences are saved locally, and are immediately (real time) sent to the control room, that by aid of CV and ML techniques, is capable to detect the hot spots,strike lines and other thermal events and send an alarm in the critical cases. Related to image analysis, that is the main topic of this thesis,

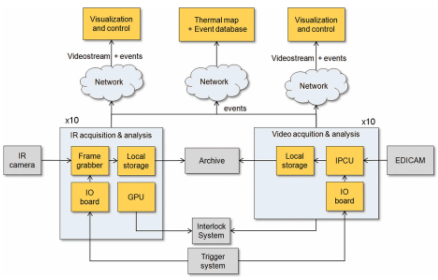
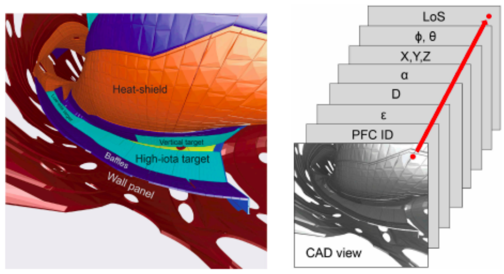


Figure 2.1: W7-X Image system overview. Extracted by [1]

the system is capable to detect and classify hot spots and strike-line shape in real time. It uses the temperature and position to do this task, and in the critical

cases is able to emit alarms.
The images are stored in a scene model that in addition to images collected by the EDICAM generates another informative images with data such as emissivity ϵ, Distance to the target object or 3-D world Cartesian coordinates. This data is stored in HDF5 files in the following manner:



(a) Field of view segmentation. (b) Multi-dimensional image.

Figure 2.2: W7-X Image Scene Model. Extracted from [1]

For the purpose of detect hot spots, the system uses a temperature threshold pixel by pixel in order to get regions of interest. Then this regions are clustered with the help of the connected components algorithm and they are tracked frame by frame. To classify this detections, it should be noted that there are different ways to lead to hot spots. For instance power overload,dust particules redeposited in the subjacent material or carbon erosion. It has been proved, that the evolution of the temperature is different for each case, therefore the eq. 2.1 is a very discriminative feature used to classify the hot spots.

$$\tau(t) = \frac{T(t)}{\frac{dT}{dt}} \tag{2.1}$$

Finally the detection and segmentation of strike line can be obtained with the following expression:

$$\phi_{th} = \phi_{max} e^{-1} \tag{2.2}$$

where the characteristics of the strike lines are then projected onto the scene model.

$$norm[t, T] = \frac{\partial_t T(t)}{T(t)} \tag{2.3}$$

where $T(t)$ is the surface temperature and $\partial_t T(t)$ is the partial derivative respect time, during the rise and decay phase. A threshold is used in order to detect the surface areas.

In addition to this, it was exposed an iterative algorithm to estimate the heat transmission coefficient α. This coefficient, as can be seen in 2, represents the relation between the heat conductivity and the thickness of the layer. Small

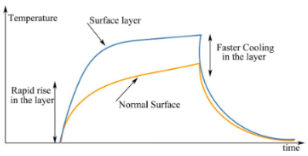


Figure 2.3: Temperature evolution of PFC. Extracted from [2]

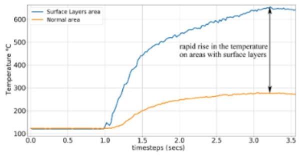


Figure 2.4: Surface temperature evolution. Extracted from [2]

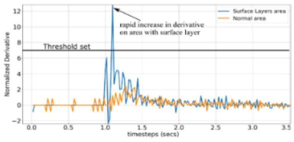


Figure 2.5: $norm_{t,T}$ and threshold. Extracted from [2]

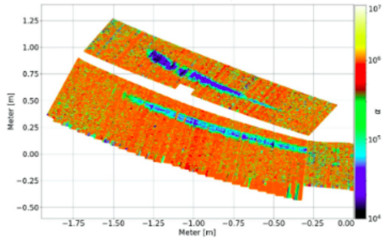


Figure 2.6: Projection of the heat flux mapped on the horizontal and vertical divertor of W7-X. Extracted from [2]

values of α represents regions of surface layers: In conclusion with IR cameras it would be interesting remark another interesting investigations with IR cameras such as [7], whereby A.puig et al. introduced the different thermal events on the plasma facing components, or [8], wherein F.Pisano et al. gives an extense overview about the essential spatial calibration and the mapping procedure.

It has been demonstrated that the majority of research carry out by the eurofusion image group has been focused on the study of IR cameras, due to the usefulness of temperature features to real time detection applications. Despite this, real cameras provide a valuable information for learn about the thermal events. This lack of specific research about PFC detections, requires rely on the work of other object recognition system. In this context, there are a massive related work such as Histogram of Oriented Gradients and Support Vector Machines for human detection[9], local binary patterns in face recognition [10] or Random Forest Classifiers in body part classification [11], or deep learning approaches like R-CNN, as proposed by Mr Chakib Belafdil et al. in Eurofusion workshop [12].

To conclude this chapter it is relevant to add that, although all related investigations are based on anomaly detection, for now is not assumable follow this approach with the EDICAMs data. Thus, the thesis will be focused on object recognition.

Chapter 3

Feature extraction techniques

Image features are essential in any CV application, such as object detection, object recognition, matching, image reconstruction, object tracking.. Related to the aim of this project, the descriptor features will be used as inputs of object recognition model. During the last 20 years, several visual descriptors has been described rely to its application. It can be defined two different image descriptors for the way in which they extract the features:

- **Local descriptor:** Describe the image patches (key points in the image) of an object. Local binary patterns [10] are one of the most well-known example of this descriptor type.
- **Global descriptor:**: Describe the image as a whole to the generalization of an object.HOG [9] is a largely used Global descriptor because of his compute ease but great performance in huge quantity of applications. [13, 14]

The best choose of Global or Local descriptor,can not be judged as isolated problem because its performance largely depends on the types of the challenge to be encountered. Generally, most object recognition systems apply either global or local features [15]. Global features are capable of generalize the whole object with a single vector, therefore their are widely used because of their computational simplicity. Local features, on the other hand, are calculated over many interest points, thereupon are more robust to occlusion and clutter, but they require specialized classification algorithms to handle the issue of different feature vector length.

The visual descriptors can be categorized in many ways,but the most important [16] are the following:

- **Color descriptors:** Color is a basic feature for image representation, and is invariant with respect to scaling, translation and rotation of an image. Due to the data generated by the EDICAM's is in gray scale, it is not possible apply this type of descriptor.

- **Texture descriptors:** Aim at capturing the general characteristics of textures in order to describe texture similarity.This descriptors could be useful with the intention of generalize the detections. In the next section will be described the *Homogeneous texture descriptor* [17] because of his extended usage and MPEG-7 implementation.

- **Shape descriptors:** Has the objective of extract the object's shape characteristics. This could be helpful in order to compare the shapes of hot spots and other detections. The descriptor *Shape Context* will be evaluated in the next section owing to great performance in other shape comparison problems.

3.1 Gradient Based Approach

Contours in the images has been one of the most largely used features in object classification as well as detection. This idea is grounded in the powerful gradients that provide the edges in the whole image, therefore they are very useful in computer vision tasks. The HOG becomes very popular because of its accuracy in pedestrian detection problems [9, 18]. This techniques how it will be explained in the next section, compute feature descriptors by way of extract histograms of oriented gradients and overlapping image regions.

3.1.1 Histogram of oriented gradients (HOG)

HOG is one of the most commonly used methods for object detection. The method was presented by N.Dalal and B.Triggs [9] and aims to evaluate well-normalitzed oriented histograms by means of generate a dense grid in the image. Objects with same shapes and appearance tends to have similar gradients, therefore the feature length can be used as a input of a classification model. HOG pass absolute pixel values over, consequently it has been proved that it is very robust and invariant against brightness, shadows and illumination changes.

The general idea of the method is the following: the local appearance of an image can be characterized properly by means of a distibution of gradients (directions) of the edges of an image. Considering this essential point, they developed a descriptor based on local gradient histograms of equally and overlapping spaced regions through sliding windows.

The implementation was described in the survey [9], the main steps are:

1. **Gamma/Colour Normalization:** It were evaluated various input pixel representations such as RGB,LAB or grayscale with gamma equalitzation.

The normalization doesn't improve considerably the performance, therefore is an optiaonal stage in the implementation.

2. **Gradient Computation:** Is one of the main step in the HOG calculation. Its purpose is detect the gradients of the images by means of the convolution with differential operators. This gradients are handy to generalize the object shapes in a image.

3. **Orientation binning:** In this part, the histograms for each cell are computed.This step introduces nonlinearities in the descriptor which are very useful in the generalization task of the descriptor.The gradient magnitude is computed and stored in a histogram bin which determines its orientation. A popular set-up of this algorithm is select 9 bins with a unsigned phase (from 0 to 180^o), whereby each histogram bin has a spread of 20. Every pixel in the cell is allocated following the next rule:

$$\begin{cases} w_k(x,y) = max(0, 1 - \frac{\theta(x,y)-\theta_k}{\delta\theta}) \\ h(k) = \sum_{(x,y)\in C} w_k(x,y)g(x,y) \end{cases} \tag{3.1}$$

   In consequence, HOG tries to assign each magnitude gradient pixel to the two nearest bins, in order to achieve a more robust descriptor than ones that only use 1 bin for each pixel.

4. **Descriptor blocks:** To be compact and robust againtst the illumination contrast, the gradient has to be locally normalized. To do this task, the cells are grouped into larger pixel regions called blocks. These block overlap with the surrounding blocks, therefore the orientation distribution of each cell is used more than once. As mentioned above, the R-HOG blocks are the most commonly used and the optimal parameters found was 2x2 cell blocks of 8 x 8 pixels cells. In addition, a Gaussian spatial window can be applied to every block with the objective of eliminates the weight of each pixel around the edge.

5. **Block normalization:** It is the final step of the HOG algorithm. It is very valuable in the robustness capability against illumination changes.

3.2 Local binary patterns (LBP)

The original local binary pattern (LBP) was introduced bt Ojala et al [10] as a non-parametric texture patterns descriptors.
LBP has a lot of relevant properties for image descriptor such as illumination changes tolerance or computational simplicity.
The basic idea behind LBP is to encode the information of the 3x3 neighbourhood of a image patch. It can be seen as a 3x3 filter, in which the center pixel is

considered to be the threshold, and each neighbouring pixel is compared against the center point in order to produce a 8-bit binary string. Then, the binary string is finally converted to decimal label according to assigned weights. In this context, using a 3x3 neighbourhood, a maximum of 256 (2^8) different textures can be described.
The result of LBP, given a pixel can be expressed as follows:

$$LBP_{P,R}(x_c, y_c) = \sum_{P=0}^{P-1} s(i_p - i_c)2^P \tag{3.2}$$

where i_p are the P surrounding pixels to central pixel (i_c). Therefore, the function s(x) can be defined:

$$s(x) = \begin{cases} 1, & if \quad x \geq 0 \\ 0, & if \quad x < 0 \end{cases} \tag{3.3}$$

The histogram of LBP encoding of a zone is suitable to be used as a image descriptor.

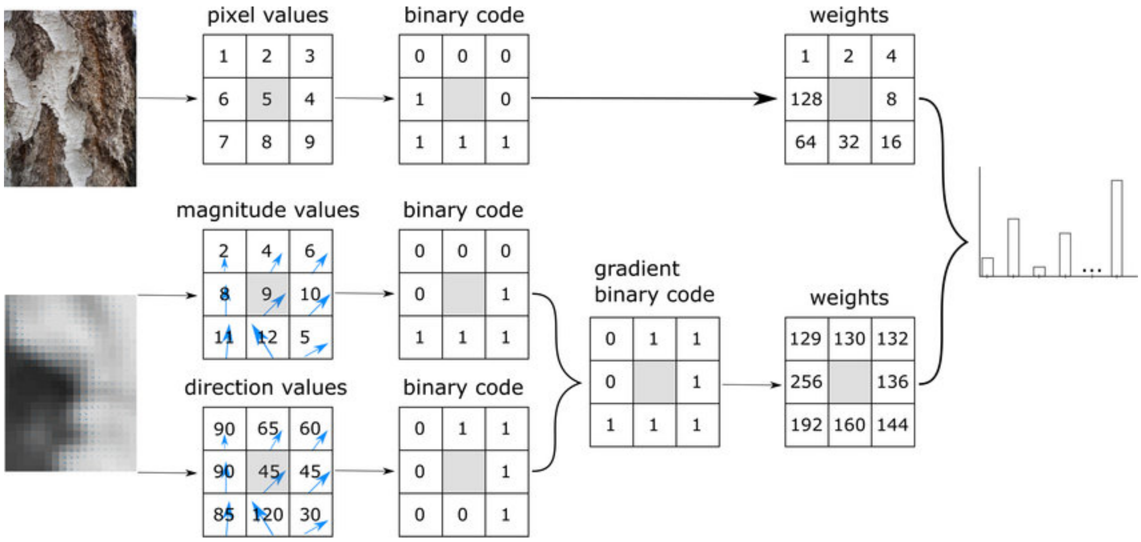


Figure 3.1: LBP operationing

3.3 Key Points Approach:Sift

This type of methods aims to extract the local image features scattered at some distributed points (key points). Then the classifier makes a final decision bear in mind the classification of each key point. The basic idea of this approach is that the key points detected are allocated in stable and robust regions of the image,

in consequence are very informative in terms of local image substance.
The precisness of classify this overall features, lies in the robustness and accuracy of each key point for each class and the additional informacion in the close pixels to the key point. One of the most used techniques of this perspective are Förstner [19], Laplacian [20] or Difference of Gaussians [21].
With the combination of this methods above-mentioned, it is possible compute a feature vectors over local image key points. The essential advantage of this focus against other thechniques such as global descriptors, is the big generalization of the image using samples of lower dimension than the total number of image pixels and thereupon accelerating the classification task.
However, these methods have some limitations just as the diferent number of key points detected, but it can be solved this issue with the addition of some dimensionality reduction method such as PCA [22] or a clustering algorithm, for example K-Means [23].

3.3.1 SIFT

The SIFT method [3] is widely used in object recognition systems. The features extracted, are invariant to image rotation and scaling, consequently it is very useful in variable environments, because they can be generalized regardless of the object orientation or object view. In addition, it has been proved that the features are very robust against distorsion, additional noise and illumination variance. David Lowe break down the method in 4 parts:

1. **Scale-space extrema detection:** In this step, the algorithm searches over all scales and image locations. It is implemented by means of the DoG function with the aim of pinpoint the potential interest points which are invariant to scale and orientation.

2. **Key point localization:** Key points are identified as local maxima or minima of the DoG images through scales. Each pixel in the DoG images is compared with its 8 neighbours at the same scale and with the 9 neighbours of the contiguos scales. If the pixel becomes as a local maximum or minimum it will be selected as candidate key point, as we can see in the Figure 3.3
   Finally the interpolation of the nearby data to the key point candidate is applied in order to determine its position more efficiently,so the key points with low contrast are removed and responses along edges eliminated. Once this processing task is made, it is assigned an orientation to key point.

3. **Orientation assignment:** With the aim to determine the key point orientation, a HOG is calculated in the surrounding key point data.

4. **Keypoint descriptor**: Once the orientation of a key point is got, it can be calculated the feature descriptor as an array of orientation histograms

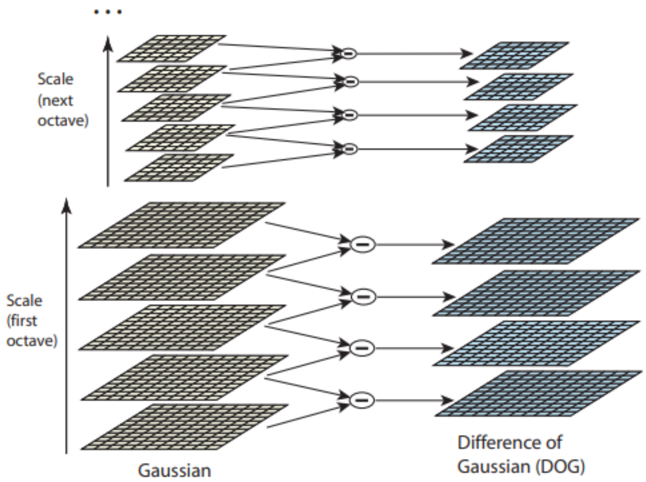


Figure 3.2: The DOG method in each octave

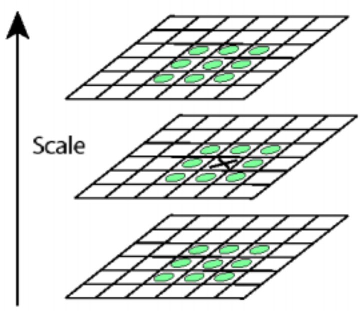


Figure 3.3: Local minimum or maximum detection. The pixel X is compared with its 26 adjacent pixels through each belonging and adjacent scale

on 4x4 pixel neighbourhoods. The parameters of the elements involved in

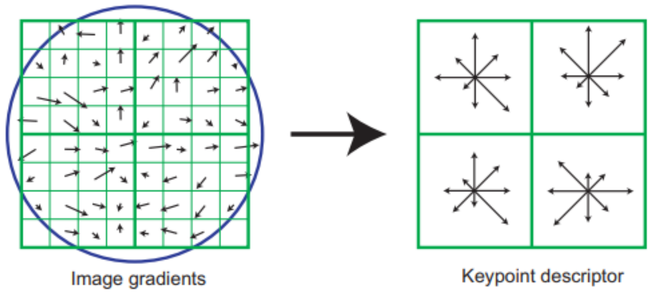


Figure 3.4: At left the gradients of analyzed image. At right the key point descriptor exposed in [3]

the SIFT method are the following:

- 8 bins by histogram
- 4 histograms around the key point

This entails a feature vector with a dimensionality of 4x4x8=128. This vector is also normalized to achieve robustness respect to changes in illumination.

Chapter 4

Feature selection techniques

4.1 Filter Based methods

4.1.1 Pearson Correlation

4.1.2 Chi-Squared

4.2 Wrapper-based methods

4.2.1 Recursive Feature Elimination

4.3 Embedded methods

4.3.1 Lasso

4.3.2 Tree-Based

Chapter 5

Classifiers

It can be defined a classifier as any algorithm that categorize data with a label. This it is possible with the help of the feature vectors that aims to be distinguished among different classes and as similar as possible regard themselves when they belong to the same class.
Typically classification algorithms include two steps, training and test:

- **Training phase**: Each feature vector is labeled by its class, and algorithm try to adjust its parameters by means of reduce an error metric such as Log Loss [24], cross-entropy [25] or Root Means Squared Error (RMSE) [26]. Once the classifier is trained, it will be available to make a classifications.
- **Test phase**: At this step, the labels of feature vectors (not seen before by classifier) in dataset are predicted with the objective of assess the final performance. This can be obtained through metrics such as F1-Score,Accuracy or Recall [27]

Classification algorithms are an example of supervised learning, which means that they need a labeling process of the feature vectors to classify. This process in computer vision is very critical, since generally there are more representatives of the object to detect in comparison with the other classes. This could lead to the undesired over-fitting [28], which means that classifier only generalize the training data, but its performance with unknown data is very poor. In spite of this, there are some techniques to deal with this issues, that will be explained in the next sections as well as the classifiers used in this thesis.

5.1 Machine learning classifiers

5.1.1 Supported vector machines (SVM)

As it mentioned before, classification of the given data is a common task in machine learning. SVM is one of the successful supervised learning methods that have strong theoretical foundations in solving classification problems. A well parameter estimation of SVM decision function would be capable of classify unseen samples accurately. One of the main reasons of the achievement of the SVM is its generalization capability combined with simple linear classification basis. The basic idea behind the SVM [29] is that given a set of a d-dimensional vectors, tries to separate them with the aid of R^d dimensional hyper plane. There are infinite hyperplanes to separate the data, but SVM intend to find the most optimal hyper-plane possible. In this context appears the definition of margin which is the distance between the nearest samples on both sides of the hyperplane. The SVM algorithm was designed with the objective of find the hyperplane with the largest margin between the two classes. If such hyperplane exists, it is called maximum-margin as we can see in the Figure 5.2.

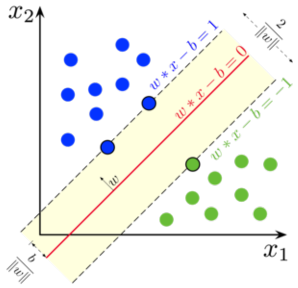


Figure 5.1: Maximum-margin hyperplane and margins for a SVM trained with two classes samples. The samples contained in the margin are called support vectors

In the following sections, it can be described the basic math theory of different SVM cases.

5.1.1.1 Linear Support Vector Machines: Hard-Margin case

Firstly, it will be seen the case when data is clearly separable linearly. It wants to achieve a classification function F:$\mathbb{R}^d \mapsto -1, 1$, which means that given a feature vector d-dimensional, the algorithm is capable to decide its class.
SVM uses the sign of a linear transformation as classifier function: $F = sign(f(x))$. $f(x)$ could be defined as a funtion such that $\mathbb{R}^d \mapsto \mathbb{R}$. The decision function essentially consists in hyper-plane able to divide the space in two classes. The hyper-plane can be define as follows:

$$\begin{gathered} f(x) = \mathbf{w^t x} + b \\ f(x) = w_0 * x_0 + w_1 * x_1 + ...w_{d-1} * x_{d-1} + b \end{gathered} \tag{5.1}$$

Where $\mathbf{x}$ represents the feature vector,$textbfw$ is the weighting vector and b is the bias.

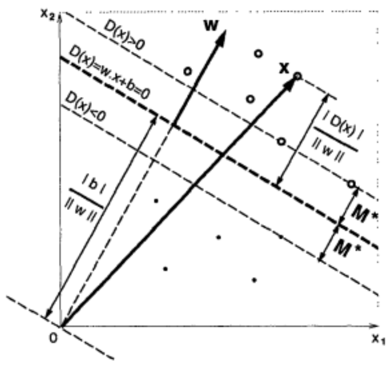


Figure 5.2: Maximum-margin linear decision function $D(\mathbf{x}) = \mathbf{w}{\cdot}\mathbf{x} + b$

5.1.2 Decision tree classifiers

Till now, it has been taken account pattern recognition algorithms based on feature vectors of real-valued and discrete-valued numbers such as SVM or K-means, and in all cases there has been a natural measure of distance between such vectors [30]. But there are very powerful algorithms, that aims to classify feature vectors using nominal data, including for example descriptions. Although feature vectors normally do not use nominal data and this type of classifier is very simple compared to other advance techniques such as SVM, it has been proved [31, ?] that they can achieve fine results in object classification. Especially remarkable is the use of decision trees based classifiers [30] with the use of the machine learning ensemble meta-algorithms, for instance Bootstrap aggregating (or bagging) or Boosting, that it will be defined in the following subsection.
Turning now to the decision tree classifiers, it is clearly understandable split the data in classes with the aid of sequence of questions, in which, the next question asked depends on the answer to the current question. The questions can be asked on a binary way: *"yes/no"* or *"true/false"* or in computer vision cases *"<threshold/>threshold"*. The sequence of questions is shown in a simply tree, where the *root node* is displayed at top, connected by subsequent *branches* to other nodes. These are connected until up to terminal or *leaf nodes* as it can be seen in the Fig. 5.3.
The decision tree algorithm can be summarised in the following steps:

- Generate the questions to split the data. The questions observe weather or not one feature exceeds a threshold.

- Compute a metric or impurity criterion to determine the information before one question is accomplished. The most common are:

$$Gini = 1 - \sum_{n=1}^{Nc} P(i) \tag{5.2}$$

$$Entropy = -\sum_{n=1}^{Nc} P(i) log_2 P(i) \tag{5.3}$$

Where Nc=Class and P(i)=$\frac{k_i}{n}$ is the quotient between the number of vectors of class i and the total number of vectors.

- Select question that minimizes the impurity criterion: $Question = argmax_i(I(i)' - I)$ where I' is the impurity after question and I the actual value of impurity.
- Split the data with the best question
- Repeat steps 2-4 until a stop criteria is accomplished

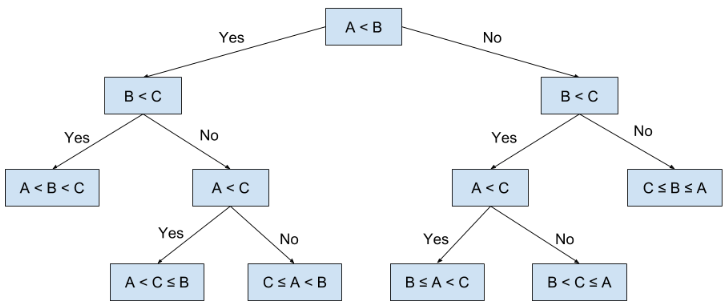


Figure 5.3: Maximum-margin linear decision function $D(\mathbf{x}) = \mathbf{w}{\cdot}\mathbf{x} + b$

Finally, one of the main aspects of decision trees is to establish when a node should be declared a leaf,in other words, when the algorithm should stop the splitting task. There are 4 approaches to lead with this issue:

1. **Threshold in impurity:** If the impurity loss of the best split is lower than a threshold it is fulfilled the stopping criteria. This perspective sometimes generates unbalanced tress.
2. **Minimum Description Length (MDL):** A criterion function is minimized:

$$\alpha \cdot size + \sum_{Nleaf} i(N) \tag{5.4}$$

Where size represents the number of nodes and α is a tuning positive parameter that controls the relation between tree size and willingness to fit the data. At this context large values of α can result in smaller trees.

3. **Validation techniques:** By means of cross-validation techniques is possible evaluate which node size provides a better performance.

4. Stop when a node present less than γ % of vectors.

5.1.2.1 Bagged decision trees: Random Forest Classifier

Random forest is one of the most popular bagging techniques in machine learning. It uses a % of training data and a $d' < d$ features in order to obtain weak decision trees. Then this classifiers are combined to obtain a final decision by voting.

Algorithm 1: Random Forest

Input : X_{Train};
$X_{Validation}$;
X_{Test};
M;
Output: Trees,$Validation_{Prediction}$,$Test_{Prediction}$
for $i = 1 : M$ **do**
  Select random subset of $d' < d$ features to train a and validate a new $Tree(i)$ with $X_{Train}(:, d')$ and $X_{Validation}(:, d')$;
  $Validation_{Prediction} \leftarrow$ prediction of $X_{Validation}(:, d')$ make use of $Tree(i)$;
  $Test_{Prediction} \leftarrow$ prediction of $X_{Test}(:, d')$ make use of $Tree(i)$;
end
Pooling predictions of each tree and decide the most probable class.

5.1.2.2 Boosted trees: Adaboost classifier

AdaBoost or Adaptive Boosting [32], is a ensemble technique which is widely used to improve the performance of lots of ML algorithms. Boosting is based in the construction of successive classifiers from selected samples of training data that take into consideration the errors of the earlier classifiers. This tools are very useful because making use of weak learnes such as decission trees, they can achieve successful outcomes in comparision with more advanced classification algorithms, for example, SVM [33].
AdaBoost [32] is intended only for binary classifiers, but there are some extensions of the AdaBoost that allows the classifier to make multi-class predictions. Either way, all of this extensions are derived by the Forward stagewise additive modeling. The Adaboost algorithm can be defined as follows:

Algorithm 2: Train AdaBoost

Input : X_{Train};
$X_{Validation}$;
M;
Output: Trees
PDF(i) = 1/N;
for $i = 1 : M$ **do**
    Train weak learner Tree(i) using samples from X_{Train} according to the distribution PDF;
    E ← Training error computed using;
    $\alpha \leftarrow \frac{1}{2} ln[(1 - E)/E]$;
    for *j=1:length*(X_{Train}) **do**
        if X_{Train} *classified correctly by Tree(k)* **then**
            PDF(j) = PDF(j) $e^{-\alpha}$;
        else
            PDF(j) = PDF(j) e^{α};
        end
    end
end

5.1.2.3 Xgboost classifier

The gradient boosting algorithm combines the gradient descendent methods and boosting techniques described in the section 5.1.2.2.

The main idea of XGBoost [34], is minimize the objective function that is in essence a loss function with regularization. The hyperparameters involved in this function are estimated by the way of gradient descend algorithm.
Once the error is minimized, it is possible build the next learner take advantage of the q value. q represents the quality of the tree due to the impurity in the weak learners classifiers.

Chapter 6

Resampling methods

Imbalanced data in Machine Learning represents the problem with classification task where the classes are not represented equally.This can lead to the scenario,in which the classifier performs a perfect fitting with the class more represented, but does not the same behaviour with the class with less samples.
This issue can be extrapolated to the situation with thermal events detection in W7-X. There are a largest quantity of detections of normal over-heating situations than hot spots, which may lead to classifiers that only classify accurately the first class, which would not be worthwhile. In this context appears the resampling techniques, which are addressed to equalize the samples in the datasets. In this thesis will be studied two approaches: over-sampling and under-sampling.

6.1 Over-Sampling: SMOTE

Synthetic Minority Over-sampling Techniques, better known as SMOTE is a over-sampling technique introduced by Nitesh V.Chawla et al. in [35]. The appear obvious focus with this approach may be duplicate random examples in the minority class. Although this technique would achieve the equalization of the classes, it doesn't add any new information to the model, therefore its usefulness would be limited. In contrast, appears SMOTE, which generates new synthetic samples draw from the availables in the dataset.

Its operation is very clear:

1. Select random sample of the minority class.
2. Select k of the nearest neighbors (this k is chose by the user, but typically its value is 5).
3. Compute a point between the sample selected and its k-nearest neighbourhood.

6.2 Under-sampling: Near-Miss

NearMiss [36] is an under-sampling method. It has the objective of balancing the classes within the dataset by randomly removing majority class representatives. When the samples are very close between them, the algorithm aims to eliminate the vectors in order to increase the spaces between the two classes. With a view to prevent the information loss (suffered in most under-sampling methods), near-neighbour methods are used. The algorithm can be summarized as follows:

1. The method compute the distances between all the majority class samples with the minority class representatives.
2. Then, the n majority class vectors with the smallest distances to the samples in the minority class samples are selected.
3. Remove the samples.

Chapter 7

Experiments and results

With the aim of evaluate each technique previously mentioned, it was necessary to process the sequences provided by IPP with the help of the commonly used libraries in python language. The procedure was as follows:

1. **Dataset Generation:**In this step, for each detection (sample), HOG and LBP descriptor were implemented as well as other space characteristics just as spot size,wide,length or intensity histogram. Once the features are extracted, it is possible label every plasma event detection.
2. **Feature descriptors evaluation:** In this section, all datasets generated will be trained by one or more classifiers in order to compare their performance. Then the two feature descriptors will be joined with the aim of camparing their performance against separately [37]. This descriptors only take advantage of spatial characteristics. Temporal characteristics study, will be evaluated by Lluis Cabello, the another student that is researching about the thermal events in W7-X.
3. **Classifiers:** At this stage, SVM, Random Forest Classifier,Adaboost, XGboost and CNN will be evaluated with the objective of establishing conclusion about which approach is better: sophisticated machine learning algorithms, decision tree based models or deep learning models.
4. **Imbalanced data techniques**: At this point, the resampling techniques: resampling and oversampling as a way to balanced the data and increase the performance of the previously trained classifiers.

7.1 Train procedure and assessment metrics

Fourteen EDICAM sequence were available that recorded thermal events. Each sequence has the following properties:

- Resolution: The edicams provides a 1280 x 1024 frames

- Framerate: 100 FPS
- Between 700-1300 frames in each sequence. On account of its framerate, each sequences only contains between 7 and 13 seconds, but the high resolution of the camera generates files with more than 1 GB.

And it can be observed the subsequently events: At left image it can be seen the

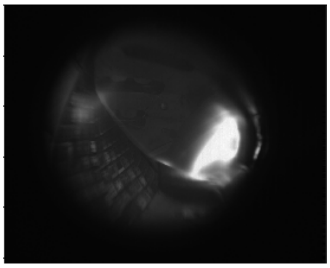


Figure 7.1: Normal behaviour

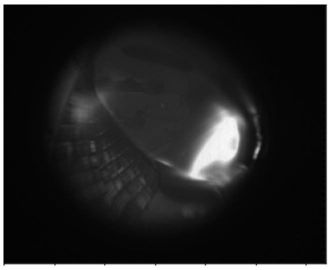


Figure 7.2: Hot spot

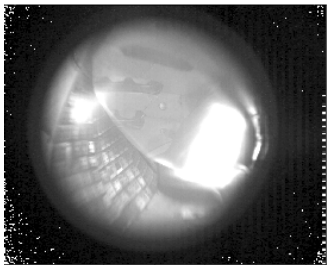


Figure 7.3: Hot spot image processing

reactor in the steady state. The other two represents the hot spot detected. It becomes clear the challenge of detect hot spots, since it is very difficult to detect without an appropriate image processing.

Once the sequences was labeled, it was generated 191718 samples of data, that were splitted in 85% (161043) for training () and 15% (30675) for test. It should be noted that the sequences used for test are independent respect to the training, since it was verified that use the same sequences for train and test produced the unwanted overfitting. The problem was that the three resulting classes are highly imbalanced:

- **Normal behaviour:** Represents the regions where is expected to have overheating because of the nuclear process. This is the class dominant with the almost 75% of samples in the train set.
- **Hot spot:** Represents the regions where is very dangerous have excessive heating. Pose about 16% of the training data.
- **Other anomalies:** Depicts the detections that can not be classified easily without the help of subject matter expert. It constitutes the remain percentage of the data.

This disproportion betweeen classes force to choose a correct evaluation metrics in order to not misread the results.

7.1.1 Cross-Validation

Cross validation [38], is a method used to evaluate the statistic analysis results and guarantee that they are independent of the split selected. In this project were used the K-Fold cross validation because its big robustness in problems with few samples.
Its operation is straightforward but powerful: it divides the training dataset into K equally parts. Subsequently, it will be done K classifications, using the K-1 subsets as train and 1 subset as validation in each classification. Then the subset used for validation change its fold in each iteration, as can be seen in the following picture: Once they have been realized the K training, the performance metrics

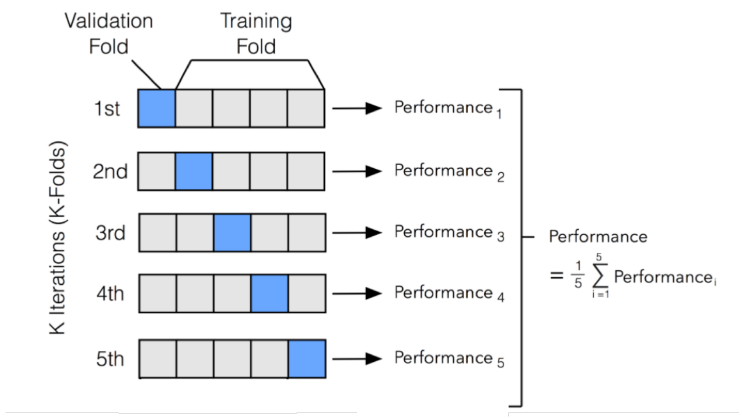


Figure 7.4: K-Fold operation

are averaged.
The K selected in this project unless it were specified the opposite, will be 5.

7.1.2 Assessment metrics

Metrics to evaluate a classifier system are necessary in every ML problem. It will be utilized the three most widely known: precision,recall and F1-Score.
Before introducing them, it is necessary explain the next concepts:

- **True Positives:** In the case of the system developed, there is the number of hot spots correctly identified.
- **False Positives:** Represents the number of non hot spot detections classified as hot spots.
- **True Negatives:** They are the non-hot spot detections labeled properly.

- **False Negatives:** They are the hot spots classified as normal behaviour regions. they are the most dangerous element for the safety of the reactor.

Precision represents the capability of the system to detect accurately the hot spots among all samples classified as hot spots. A high precision would indicate the lack of false alarms. It can be computed as follows:

$$Precision = \frac{TP}{TP + FP} \tag{7.1}$$

On the other hand, recall is the ability to classify the hot spots correctly. A low recall would show that a huge number of hot spots are not being detected, which would not be safe for the reactor.
Therefore, our system must prioritize the recall against the precision. The recall can be calculated in the following manner:

$$Recall = \frac{TP}{TP + FN} \tag{7.2}$$

Finally it will be used the F1-Score, which is a helpful metric that links precision with recall. It can be estimated on the following way:

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{7.3}$$

Finally, it is interesting the use of the area under the curve[39], with the objective of analyze the classifier behaviour at various threshold settings.

7.2 Image descriptors and feature extraction

Firstly, once the detections are stored in a JSON file, it were implemented the feature extractor system to characterize every sample detected. It provides the following characteristics for each recognition:

- Binarized area
- Width
- Height
- Form factor($width/height$)
- Fill factor ($area/(width \cdot height)$)
- Centroid X coordinate
- centroid Y coordinate

- mean detection
- mean overall frame
- intensity histogram
- feature descriptor histogram
- mean ring around hot spot
- ring around hot spot histogram

A priori, feature descriptor histogram should be the most relevant feature, therefore they were implemented two approaches with their respective parameters:

7.3 Gradient global approach(HOG)

HOG was the global and gradient based approach chose owing to big performance in shape detection research. It had the next parameters:

- **Size window:** 64 x 64
- **Block size:**16 x 16
- **Block Stride:**8 x 8
- **Cell size:**8 x 8
- **Orientation bins:** 9
- **Levels:** 64

7.4 Local key point approach(LBP)

It was the local descriptor chose because its simplicity and non time-consuming operation. It was implemented by Marti and had the following parameters:

- **Neighbours:**8
- **Circle radius:**2
- **Bins:**256

Once, generated the dataset the validation AUC over K-Folds was the following:

7.5 Classifiers

In order to analyze the classifiers behaviour, they were implemented the classifiers described in chapter 5. With the aim of increase their performance,a grid search was used that made all possible combinations of hyperparameters with the objective of selecting the best option. It were combined iteratively all parameters in order to find the optimal combination.This results were validated with K-Fold cross-validation. The hyper parameters were the following for each classifier:

- **SVM:** C parameter that defines the margin of the SVM, γ that defines the kernel coefficient and the kernel type transformation (linear, rbf..). After the grid search, it was chosen a rbf kernel, a value of 10 for the C and 0.01 for the gamma.
- **RFC:** Class weight that is very useful to deal with the imbalanced dataset,the number of decision trees, the max depth of a node, and the impurity criterion. Post grid search it was founded the next values: Entropy criterion, max tree depth of 9, and 600 trees).
- **Adaboost:** Adaboost uses a decision tree as a weak learner, therefore has the same hyperparameters than RFC
- **XGboost:** XGboost is a tree-based model also, therefore has the same parameters as RFC and Adaboost except for learning rate used in the computation of gradient descent. The learning rate chosen was 0.01.

7.6 Imbalanced data techniques

The imbalanced data techniques used in these experiments had a basic algorithms behinds therefore it didn't need for hyperparameter tuning. It were used the 3 nearest neighbourhoods in the both oversampling and undersampling methods.

7.7 Final results

The following table summarizes the experiments with 2 classes and only with the features extracted in order to compare the quality of the color descriptors:

	Train			Validation			Test		
Method	**P**	**R**	F_1	**P**	**R**	F_1	**P**	**R**	F_1
LBP+SVM	1	1	1	0.996	0.991	0.993	0.91	0.95	0.93
LBP+RFC	**1**	**1**	**1**	**0.99**	**0.97**	**0.98**	**0.99**	**0.97**	**0.98**
LBP+ADA	1	1	1	0.998	0.996	0.997	0.96	0.98	0.97
LBP+XGB	1	1	1	0.998	0.996	0.997	0.93	0.97	0.95
HOG+SVM	0.999	0.999	0.999	0.98	0.934	0.956	0.903	0.928	0.915
HOG+RFC	**0.999**	**0.999**	**0.999**	**0.999**	**0.958**	**0.978**	**0.974**	**0.998**	**0.986**
HOG+ADA	0.999	0.999	0.999	0.999	0.943	0.97	0.968	0.983	0.975
HOG+XGB	1	1	1	0.998	0.992	0.995	0.917	0.971	0.942

Table 7.1: Comparison of the color descriptors with 2 classes approach

As can be seen, it had obtained unbelievably good results when anomaly class is added to dominant class, therefore it had no sense to continue the experiments with the two classes approach. Henceforth, the testing realized were with the three classes mentioned above. Another interesting result observed was the proper performance achieve with RFC against the other classifiers and its different execution time response:

Method	***Exection time (hours)***
LBP+SVM	16:33
LBP+RFC	**0:10**
LBP+ADA	10:08
LBP+XGB	4:27
HOG+SVM	22:45
HOG+RFC	0:38
HOG+ADA	13:35
HOG+XGB	6:15

Table 7.2: Classifiers execution time

Therefore, the subsequently experiments were realized with the LBP feature descriptor in combination with RFC and XGBoost.

Once the feature descriptors and classifiers were compared with two classes, the problem become more tough in the multiclass classification. The matrix confusion obtained were the following:

$$RFC_{normal} = \begin{bmatrix} 19607 & 1940 & 6712 \\ 5 & 48 & 1108 \\ 292 & 8 & 955 \end{bmatrix} \tag{7.4}$$

$$RFC_{SMOTE} = \begin{bmatrix} 27499 & 1 & 759 \\ 52 & 1109 & 0 \\ 691 & 14 & 550 \end{bmatrix} \tag{7.5}$$

$$XGBoost_{Normal} = \begin{bmatrix} 15671 & 6207 & 6381 \\ 2 & 1150 & 9 \\ 350 & 81 & 824 \end{bmatrix} \tag{7.6}$$

$$XGBoost_{SMOTE} = \begin{bmatrix} 26854 & 432 & 973 \\ 1 & 1154 & 6 \\ 630 & 38 & 587 \end{bmatrix} \tag{7.7}$$

Only with simply exploration it has been demonstrated that resampling methods significantly improves the performance of classifier. In terms of assessment metrics class by class the results were the next:

	Normal overheating			**Hot spot**			**Other anomaly**		
Method	**P**	**R**	F_1	**P**	**R**	F_1	**P**	**R**	F_1
LBP+RFC	0.985	0.693	0.8141	0.024	0.041	0.03	0.11	0.76	0.192
LBP+RFC+SMOTE	0.973	0.973	0.973	0.986	0.98	0.983	0.42	0.438	0.428
XGB+RFC	0.978	0.55	0.704	0.154	0.99	0.266	0.114	0.656	0.194
XGB+RFC+SMOTE	0.977	0.95	0.963	0.71	0.993	0.828	0.375	0.467	0.415

Table 7.3: Comparison of the SMOTE performance between RFC and HOG

As can be seen oversampling techniques offers an hopeful results in the detection of hot spots as well as normal over-heating regions. The only snag is the blunder classifying other anomalies.

Finally it is important to highlight that the code written to realize those experiments can be found in the following link. Thesis-Code

Chapter 8

Conclusions

This thesis had reported an overview about the different strategies carried out in the object recognition task. The lack of researching about thermal events detection with the EDICAMs leads to experiment with the best-known computer vision systems in order to solve the hot spot threat in a nuclear fusion reactor as well as acquire knowdlege about the disruptions during the quasi steady-state operation.

First, the different approaches of image descriptors were proposed and investigated with the aim of analyze which adjust more to the images of over-heating situations inside the stellarator recollected by the EDICAMs. It was focused on LBP, HOG and SIFT because its success in other CV applications.

Then, following the typical scheme of the object recognition systems, it were investigated the best-known classifiers. Nowadays the decision tree based algorithms popularity is increasing because its simplicity and non time-consuming. Therefore three different approaches of this type of algorithms were explored in addition with SVM, because of this extended use in artificial intelligence.

The last part of the study was the evaluation of the techniques to deal with imbalanced data sets. As mentioned in introduction, the initial detection system, just take into account the brightness regions of the image, therefore there are a lot of FP. Once labeled, the percentage of hot spots were much less than the normal overheating regions. Thus, resampling techniques become essential to achieve best results.

Finally, the experiments have shown that the feature extraction techniques as well as classifiers chosen were on the right way. In spite of lack of experts validation, the results are hopeful and defines the way to follow in the following investigations.

Chapter 9

Future Work

This thesis has focused on the feature extraction and classification of thermal events. Following this point,there is a still of work to do, such as implement other feature descriptors used in other computer vision applications (SURF,FAST, BRIEF, Haralick ..) or use most sophisticated deep learning techniques just as fully connected layers to classify the features extracted, or Fast R-CNN with the aim of classify the images directly.

In addition to this there are other points to be investigated that have not been covered in this thesis. For instance, the improvement of the detection system with the aim of reduce the FP, the temporal feature extraction or other trackers approaches instead of the current centroid focus.

Finally it would be interesting to change the approach of this project and research about the anomaly detection rather than object recognition. This is not a easy problem, because the EDICAMs collect the thermal events more slowly than IR, but try to detect the hot spots directly would increase the interest of the survey because of its online properties.

Bibliography

[1] A. P. Sitjes, M. Jakubowski, A. Ali, P. Drewelow, V. Moncada, F. Pisano, T. T. Ngo, B. Cannas, J. M. Travere, G. Kocsis, T. Szepesi, T. Szabolics, and W.-X. Team, "Wendelstein 7-x near real-time image diagnostic system for plasma-facing components protection," *Fusion Science and Technology*, vol. 74, no. 1-2, pp. 116–124, 2018. [Online]. Available: https://doi.org/10.1080/15361055.2017.1396860

[2] A. Ali, H. Niemann, M. Jakubowski, T. S. Pedersen, R. Neu, Y. Corre, P. Drewelow, A. P. Sitjes, G. Wurden, F. Pisano, B. Cannas, Y. Gao, and M. Ślęczka, "Initial results from the hotspot detection scheme for protection of plasma facing components in wendelstein 7-x," *Nuclear Materials and Energy*, vol. 19, pp. 335 – 339, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S2352179118301698

[3] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, p. 91–110, 2004.

[4] A. Ledda, "Mathematical morphology in image processing," Ph.D. dissertation, 01 2007.

[5] L. He, X. Ren, Q. Gao, X. Zhao, B. Yao, and Y. Chao, "The connected-component labeling problem: A review of state-of-the-art algorithms," *Pattern Recognition*, vol. 70, pp. 25 – 43, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320317301693

[6] Mau, A. Rosebrock, Venkad, Bosmart, Ngapweitham, Phong, Farah, C. Nolet, Mooj, Walid, and et al., "(faster) non-maximum suppression in python," Apr 2020. [Online]. Available: https://www.pyimagesearch.com/2015/02/16/faster-non-maximum-suppression-python/

[7] A. P. Sitjes, Y. Gao, M. Jakubowski, P. Drewelow, H. Niemann, A. Ali, V. Moncada, F. Pisano, T. Ngo, B. Cannas, and et al., "Observation of thermal events on the plasma facing components of wendelstein 7-x," *Journal of Instrumentation*, vol. 14, no. 11, p. C11002–C11002, Nov 2019. [Online]. Available: http://dx.doi.org/10.1088/1748-0221/14/11/c11002

[8] F. Pisano, B. Cannas, M. W. Jakubowski, H. Niemann, A. Puig Sitjes, and G. A. Wurden, "Towards a new image processing system at wendelstein 7-x: From spatial calibration to characterization of thermal events," *Review of*

Scientific Instruments, vol. 89, no. 12, p. 123503, 2018. [Online]. Available: https://doi.org/10.1063/1.5045560

[9] B. T. N. Dalal, "Histograms of oriented gradients for human detection," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05) - Workshops.*

[10] D. Huang, C. Shan, M. Ardabilian, Y. Wang, and L. Chen, "Local binary patterns and its application to facial image analysis: A survey," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 41, no. 6, p. 765–781, 2011.

[11] S. Handrich, A. Al-Hamadi, E. Lilienblum, and Z. Liu, "Human body-part classification using geodesic descriptors and random forests," in *2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA)*, 2017, pp. 318–321.

[12] "Eurofusion workshop on image processing and machine learning." [Online]. Available: https://indico.euro-fusion.org/event/48/?print=1&view=nicecompact

[13] O. Déniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," *Elsevier BV*, vol. 32, no. 12, p. 1598–1603, 2011.

[14] P. E. Rybski, D. Huber, D. D. Morris, and R. Hoffman, "Visual classification of coarse vehicle orientation using histogram of oriented gradients features," in *2010 IEEE Intelligent Vehicles Symposium*, 2010, pp. 921–928.

[15] D. Lisin, M. Mattar, M. Blaschko, E. Learned-Miller, and M. Benfield, "Combining local and global image features for object class recognition," *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR05) - Workshops.*

[16] R. M. Kumar, "A survey on image feature descriptors," *International Journal of Computer Science and Information Technologies*, vol. 5, p. 7668–7673, 2014.

[17] Y. M. Ro and H. K. Kang, "Hierarchical rotational invariant similarity measurement for mpeg-7 homogeneous texture descriptor," *Electronics Letters*, vol. 36, no. 15, p. 1268, 2000.

[18] N. Dalal, "Finding people in images and videos," 07 2006.

[19] D. Jiang and J. Yi, "Comparison and study of classic feature point detection algorithm," in *2012 International Conference on Computer Science and Service System*, 2012, pp. 2307–2309.

[20] "Feature detection with automatic scale selection," *International Journal of Computer Vision*, vol. 30, no. 2, Nov 1998. [Online]. Available: https://doi.org/10.1023/A:1008045108935

[21] A. Bundy and L. Wallen, "Difference of gaussians," *Catalogue of Artificial Intelligence Tools*, p. 30–30, 1984.

[22] I. T. Jolliffe and J. Cadima, "Principal component analysis: a review and recent developments," vol. 374, no. 2065, Apr 2016. [Online]. Available: https://pubmed.ncbi.nlm.nih.gov/26953178

[23] J. B. MacQueen, "Some methods for classification and analysis of multivariate observations," *Proceedings of 5-th Berkeley Symposium on Mathematical Statistics and Probability*, vol. 1, p. 281–297, 1967.

[24] V. Vovk, "The fundamental nature of the log loss function," *CoRR*, vol. abs/1502.06254, 2015. [Online]. Available: http://arxiv.org/abs/1502.06254

[25] J. Brownlee, "A gentle introduction to cross-entropy for machine learning," Dec 2019. [Online]. Available: https://machinelearningmastery.com/cross-entropy-for-machine-learning/

[26] J. Moody, "What does rmse really mean?" Sep 2019. [Online]. Available: https://towardsdatascience.com/what-does-rmse-really-mean-806b65f2e48e

[27] A. Mishra, "Metrics to evaluate your machine learning algorithm," May 2020. [Online]. Available: https://towardsdatascience.com/metrics-to-evaluate-your-machine-learning-algorithm-f10ba6e38234

[28] X. Ying, "An overview of overfitting and its solutions," *Journal of Physics: Conference Series*, vol. 1168, no. 2, 2019.

[29] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, p. 273–297, 1995.

[30] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification.* Wiley-India, 2012.

[31] M. Martínez Zarzuela, F. Díaz-Pernas, M. Antón-Rodríguez, F. Perozo, and D. González-Ortega, "Adaboost face detection on the gpu using haar-like features," 05 2011, pp. 333–342.

[32] Y. Freund and R. E. Schapire, "A decision-theoretic generalization of on-line learning and an application to boosting," *Journal of Computer and System Sciences*, vol. 55, no. 1, p. 119–139, 1997.

[33] H. Mady and S. M. S. Hilles, "Face recognition and detection using random forest and combination of lbp and hog features," in *2018 International Conference on Smart Computing and Electronic Enterprise (ICSCEE)*, 2018, pp. 1–7.

[34] T. Chen and C. Guestrin, "Xgboost," *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, Aug 2016. [Online]. Available: http://dx.doi.org/10.1145/2939672.2939785

[35] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: Synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, p. 321–357, Jun 2002. [Online]. Available: http://dx.doi.org/10.1613/jair.953

[36] J. Zhang and I. Mani, "KNN Approach to Unbalanced Data Distributions: A Case Study Involving Information Extraction," in *Proceedings of the ICML'2003 Workshop on Learning from Imbalanced Datasets*, 2003.

[37] S. Yang, X. Liao, and U. Borasy, "A pedestrian detection method based on the hog-lbp feature and gentle adaboost," *International Journal of Advancements in Computing Technology*, vol. 4, pp. 553–560, 10 2012.

[38] S. Arlot and A. Celisse, "A survey of cross-validation procedures for model selection," *Statist. Surv.*, vol. 4, pp. 40–79, 2010. [Online]. Available: https://doi.org/10.1214/09-SS054

[39] A. P. Bradley, "The use of the area under the roc curve in the evaluation of machine learning algorithms," *Pattern Recognition*, vol. 30, no. 7, pp. 1145 – 1159, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0031320396001422