



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MECÁNICA

IDENTIFICACIÓN DE DAÑOS EN PLACAS COMPUESTAS MEDIANTE
PROCESAMIENTO AUTOMÁTICO DE IMÁGENES

MEMORIA PARA OPTAR AL TÍTULO DE
INGENIERO CIVIL MECÁNICO

FELIPE IGNACIO GUILLEN PALACIOS

PROFESOR GUÍA:
VIVIANA MERUANE NARANJO

MIEMBROS DE LA COMISIÓN:
ENRIQUE LÓPEZ DROGUETT
NANCY HITSCHFELD KAHLER

Este trabajo ha sido parcialmente financiado por proyecto Fondecyt 1170535

SANTIAGO DE CHILE
2018

RESUMEN DE LA MEMORIA PARA OPTAR
AL TÍTULO DE INGENIERO CIVIL MECÁNICO
POR: FELIPE IGNACIO GUILLEN PALACIOS
FECHA: 2018
PROF. GUÍA: VIVIANA MERUANE NARANJO

IDENTIFICACIÓN DE DAÑOS EN PLACAS COMPUESTAS MEDIANTE PROCESAMIENTO AUTOMÁTICO DE IMÁGENES

Hoy en día el análisis de fallas o desgastes en estructuras es un punto de investigación que ha mantenido a varios científicos a la vanguardia con respecto a estos temas. Siempre se ha buscado optimizar el análisis de fallas, de tal forma de realizarlo en el menor tiempo y de la forma más precisa posible.

En este contexto y basado en el problema de la detección de anomalías de pacientes en mamografías, las cuales se detectan en gran parte en base a la experiencia de los doctores, ha dado pie a la investigación de metodologías de procesamiento de imágenes para evaluar el daño en estructuras compuestas.

La memoria comprende un estudio específico de imágenes tomadas a estructuras compuestas, por ejemplo, en paneles tipo sándwich, en donde se diseñarán métodos de procesamiento automático de imágenes para la detección de zonas delaminadas o dañadas.

Se desarrolla una metodología para estudiar los distintos tipos de métodos de reconocimiento de imágenes que se usan hoy en día en detección de masas en mamogramas y analizar cuáles son factibles a aplicar en reconocimiento de daños en estructuras compuestas. Luego se busca aplicar estos métodos en imágenes simuladas al azar para obtener algoritmos lo más eficientes posibles. Finalmente se aplican estos algoritmos en imágenes experimentales y se concluyen la eficacia de estos.

Se trabaja con 7 métodos de análisis de imágenes hasta el momento, Promedio Factor de Correlación el cual busca encontrar una cota la cual limita las intensidades de colores. Para ello se trabaja con una matriz de confusión para poder obtener la proporción de falsos positivos y negativos, la cual se aplica en todos los métodos de tal forma de obtener la certeza de cada uno. Junto con este método se implementa una mejora en donde se eliminan objetos o anomalías de menores tamaño ya que no se consideran como daños.

También se trabaja identificando los outliers considerando una distribución T-Student. Se trabaja con la caracterización del perímetro y del área de las anomalías obtenidas ya sea modelándolas por default al momento de ser procesadas o a través de una elipse con similar segundo momento normalizado. Por último también se trabaja con un método el cual compra la variación de gradiente entre una imagen procesado y una imagen con índices de daños y define las anomalías en base a la magnitud del gradiente.

Como principales resultados se obtuvo una Proporción de Falsos Negativos (FNR) de 0,046 con el método de Ponderación con Factor Correlación con Filtro, pero aplicando el método de área y perímetro en base a elipses se obtiene un valor mayor de FNR de 0,052 pero con menos variabilidad, con lo que lo hace un algoritmo más confiable pero menos certero.

Dedicado a la mejor madre de todas, gracias por todo!

Agradecimientos

Como todo en la vida, primero quiero agradecer a mi familia, a mi madre, por su apoyo incondicional, inquebrantable y que solo ella me pudo dar. Gracias por la paciencia no solo en la carrera sino en toda la vida y sobretodo agradecer por esas diferentes elecciones que tuviste que tomar con empoderamiento que me llevaron hasta donde estoy ahora.

Luego quiero agradecer a todos mis amigos, desde el colegio (Álvaro, Mati, Toño), la universidad (cabros de tropi, plan común, Dani, Conty, mecánica), amigos de trabajos de hoy en día y de días pasados (Amigos de C+, Rolo, Enano y otros que ya no me acuerdo pero me marcaron), ya que cada uno siempre tuvo algo que aporta tanto en mi crecimiento tanto como persona como profesional a lo largo de la vida y carrera.

En especial quiero agradecer al Sibona, Miguel, Goza, Joaco, en verdad a todo Tropi y aparte al Rolo, Conty, Dani...En verdad todos marcaron alguna diferencia...ya sea pasteando o enseñandome cosas importantes.

Quiero agradecer a mi polola, la Javi, por el apañe al último momento de la carrera, me diste el último pero no menos importante empujón para terminar todo. Tambien agradecer por los consejos que me diste durante el último año que cambio drasticamente mi vida profesional y universitaria.

Al profe de taekwondo (Alejandro) y handball (Victor) ya que desde el deporte me enseñaron grandes valores para la vida, cómo el no tener excusas y hacer siempre las cosas de mejor manera y con todas las ganas posibles.

Reiterar mis agradecimientos a mi madre, gracias por luchar junto conmigo para lograr lo que hemos logrado...

Desde ahora se empieza de nuevo...

Tabla de Contenido

1. Introducción	1
1.1. Motivación	1
1.2. Objetivos	2
1.2.1. Objetivo Principal	2
1.2.2. Objetivo Principal	2
1.3. Alcances	2
2. Antecedentes	3
2.1. Imágenes	3
2.2. Algoritmos de identificación de daños	4
2.2.1. Promedio Factor de Correlación	5
2.2.2. Filtros de menores tamaño	6
2.2.3. Distribución T-Student Outliers	6
2.2.4. Proporción entre el área y el perímetro del contorno	8
2.2.5. Proporción entre el área y el perímetro en base a elipses	8
2.2.6. Variación de gradientes	9
2.2.7. Filtro Área y Perímetro	9
2.2.8. Imágenes experimentales	10
3. Metodología	13
4. Resultados	15
4.1. Ponderación de Factor Correlación	15
4.2. Filtro de menores tamaño	16
4.3. Distribución T-Student Outliers	19
4.4. Proporción entre el área y el perímetro	20
4.5. Proporción área y perimetro modelada con elipses	22
4.6. Variación de gradientes	25
4.7. Filtro Área y Perímetro	26
4.8. Algoritmos en imágenes reales	28
Conclusión	33
Bibliografía	35

Índice de Tablas

2.1. Matriz de Confusión.	5
4.1. Rates obtenidos con Ponderación de Factor Correlación.	16
4.2. Rates obtenidos con Ponderación de Factores con filtro.	18
4.3. Rates obtenidos con Método T-Student.	20
4.4. Rates obtenidos en base a Proporción entre Área y Perímetro.	21
4.5. Rates obtenidos en base a Proporción entre Área y Perímetro modelada con elipses.	24
4.6. Rates obtenidos en base a Variación de gradientes.	25
4.7. Rates obtenidos con Ponderación de Factor Correlación con filtros de tamaño y área-perímetro.	28
4.8. Valores de FNR y FPR promedios de cada algoritmo en imágenes simuladas.	28
4.9. Valores de FNR y FPR para cada caso por cada algoritmo en imágenes experimentales.	29

Índice de Ilustraciones

2.1. Masas detectadas como anormales en un mamograma	3
2.2. Ejemplo de indicadores de daño.	4
2.3. Imágenes generadas al azar.	4
2.4. Metodología Promedio Factor Correlación	6
2.5. Metodología Promedio Factor Correlación con filtro	7
2.6. Metodología Proporción entre el área y el perímetro.	8
2.7. Metodología Proporción entre el área y el perímetro en base a elipses.	9
2.8. Metodología variación de gradiente.	10
2.9. Metodología Filtro Área y Perímetro.	10
2.10. Imágenes experimentales, caso 1.	11
2.11. Imágenes experimentales, caso 2.	11
2.12. Imágenes experimentales, caso 3.	12
4.1. a. Mapa de índices de daño, b. Daño detectado con Ponderación de Factores, c. Daño original.	15
4.2. Evolución Rates de FPR y FNR en función del Factor de Correlación	16
4.3. Histograma rate FNR con Ponderación de Factor Correlación.	17
4.4. a. Mapa de índices de daño, b. Daño detectado con Ponderación de Factores con filtro, c. Daño original.	17
4.5. Evolución Rates de FPR y FNR en función del Factor de Correlación con y sin filtros.	18
4.6. Histograma rate FNR con Ponderación de Factores con filtro.	19
4.7. a. Mapa de índices de daño, b. Daño detectado con Método T-Student, c. Daño original.	19
4.8. Histograma rate FNR con Método T-Student sin filtro.	20
4.9. Histograma rate FNR con Método T-Student con filtro.	21
4.10. a. Mapa de índices de daño, b. Daño detectado en base a Proporción entre Área y Perímetro, c. Daño original.	21
4.11. Función de relación entre área y perimetro.	22
4.12. Histograma rate FPR en base a Proporción entre Área y Perímetro.	23
4.13. Histograma rate FNR en base a Proporción entre Área y Perímetro.	23
4.14. a. Mapa de índices de daño, b. Daño detectado en base a Proporción entre Área y Perímetro modelada con elipses, c. Daño original.	24
4.15. Función de relación entre área y perimetro de elipses.	24
4.16. Histograma rate FPR en base a Proporción entre Área y Perímetro modelada con elipses.	25

4.17. Histograma rate FNR en base a Proporción entre Área y Perímetro modelada con elipses.	26
4.18. Histograma rate FPR.	26
4.19. a. Mapa de índices de daño, b. Daño detectado en base a Variación de Gradientes, c. Daño original.	27
4.20. Evolución Rates de FPR y FNR en base a Variación de gradientes.	27
4.21. a. Mapa de índices de daño, b. Daño detectado con Ponderación de Factores con filtros de tamaño y área-perímetro, c. Daño original.	28
4.22. Histograma rate FNR con Ponderación de Factor Correlación con filtros de tamaño y área-perímetro.	29
4.23. Imágenes procesadas con Ponderación de Factor Correlación para cada caso.	30
4.24. Imágenes procesadas con Ponderación de Factor Correlación con filtro para cada caso.	30
4.25. Imágenes procesadas con Método T-Student con y sin filtro para cada caso. .	31
4.26. Imágenes procesadas en base a Área y Perímetro, tanto en caso normal como modelada en base a elipses para cada caso, y en base a variación de gradientes.	31
4.27. Imágenes procesadas en base a Ponderación de Factor Correlación con filtro de tamaño y en relación a área-perímetro para cada caso.	32

Capítulo 1

Introducción

Esta propuesta aborda el problema del análisis automatizado de imágenes para evaluar el daño en estructuras compuestas. Las imágenes a analizar están relacionadas con la distribución de daños en la superficie de una estructura. En ellas se presenta una distribución de los indicadores de daños en un panel tipo sándwich con delaminación. La región delaminada aparece como una región con una intensidad más altas que sus vecinos.

La identificación de regiones delaminadas a partir de la distribución de indicadores de daño es similar al problema de la detección anormal de masas de mamografías. Aunque es una aplicación diferente, el problema es el mismo, las masas también se presentan como regiones con mayor intensidad en mamogramas.

El análisis automático de imágenes para mamografías ha sido ampliamente investigado en las últimas décadas y hay varias técnicas disponibles. Por lo tanto, proponemos investigar la extensión de las metodologías de procesamiento automático de imágenes implementadas para la detección de masas en mamogramas a la detección de regiones delaminadas en paneles tipo sándwich.

1.1. Motivación

Con la propuesta de analizar imágenes que muestran una distribución de indicadores de daños en un panel compuesto es posible estudiar distintos tipos de metodologías de procesamiento automático de imágenes para obtener una configuración óptima de estas. La presente memoria tiene por motivación generar un impacto científico en el análisis de estructura compuestas. Mediante el uso de la tecnología de hoy en día se pretende investigar y desarrollar un método que puede impactar en el estudio de daños en estructuras a largo plazo.

1.2. Objetivos

A continuación, se plantean los objetivos del trabajo de título.

1.2.1. Objetivo Principal

Investigar y desarrollar una metodología de procesamiento automático de imágenes para la detección de regiones delaminadas en paneles tipo sándwich.

1.2.2. Objetivo Principal

El objetivo general se desarrolla por medio de los siguientes objetivos específicos:

1. Análisis de metodologías de reconocimiento de imágenes aplicables al problema en estudio.
2. Implementación de distintas metodologías de procesamiento automático en imágenes creadas mediante una simulación numérica de los paneles con daños generados de forma aleatoria.
3. Análisis de sensibilidad de las metodologías a los distintos parámetros y funciones disponibles, y selección de la configuración que obtenga los mejores resultados en la detección de daño.
4. Aplicación y análisis de las metodologías y configuraciones seleccionadas en imágenes reales.

1.3. Alcances

El alcance de este trabajo consiste en el estudio y desarrollo de metodologías de procesamiento automático de imágenes. Para ello se realiza un proceso iterativo en donde se probará la eficacia de diferentes métodos en imágenes distintas buscando obtener una configuración óptima. La toma experimental de imágenes no está contemplada en la memoria, estas serán otorgadas por la Profesora Guía.

Capítulo 2

Antecedentes

A continuación, se presentan los antecedentes generales que contextualizan lo mencionado en este informe, de modo de lograr una mayor comprensión de este, además de los antecedentes importantes que justifican selección de los métodos principales con los que se trabajará a lo largo de la memoria.

2.1. Imágenes

Como se ha hablado el problema planteado viene de la detección de anomalías en mamografías. En la Figura 2.1. se observan unos ejemplos de mamogramas. En ellos se aprecia que la detección de dos anomalías.[1][4]

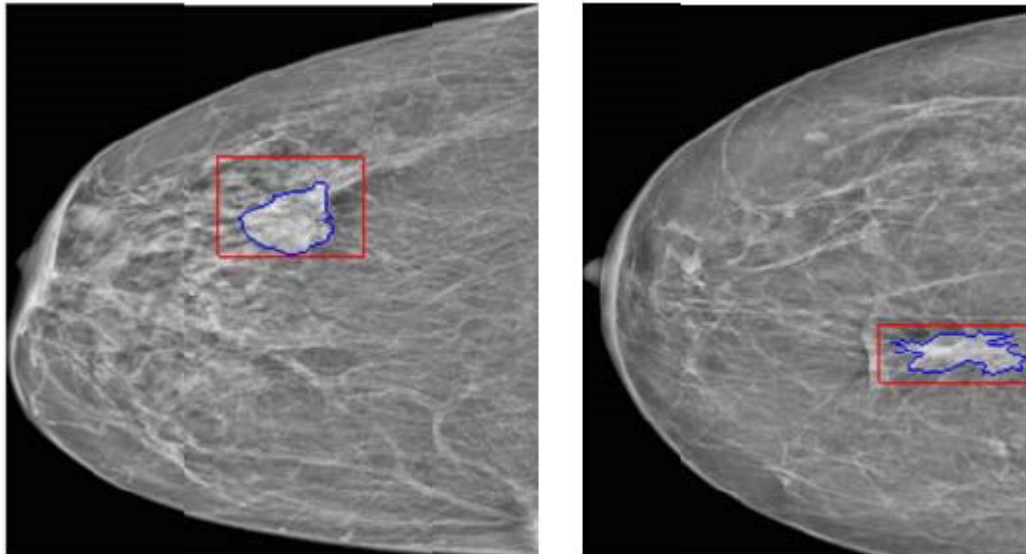


Figura 2.1: Masas detectadas como anormales en un mamograma

Las imágenes con que se probará los métodos finales muestran la distribución de los

indicadores de daños de la superficie del panel, estas son del estilo de la imagen mostrada en la Figura 2.2. Aquí se muestran los indicadores de daños en un panel tipo sandwich con delaminación.

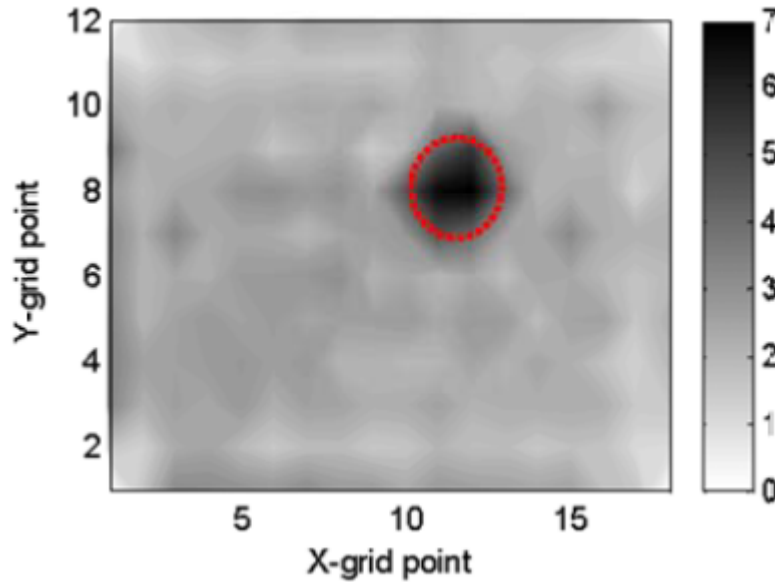


Figura 2.2: Ejemplo de indicadores de daño.

Pero las imágenes con que se harán las primeras iteraciones de los algoritmos son del estilo de las mostradas en la Figura 2.3. Estas imágenes son generadas aleatoriamente, y son otorgadas por la Profesora Guía. Se llevará a cabo una gran cantidad de pruebas con este tipo de imágenes ya que la imágenes experimentales son limitadas y toma mucho tiempo obtenerlas.[5][7]

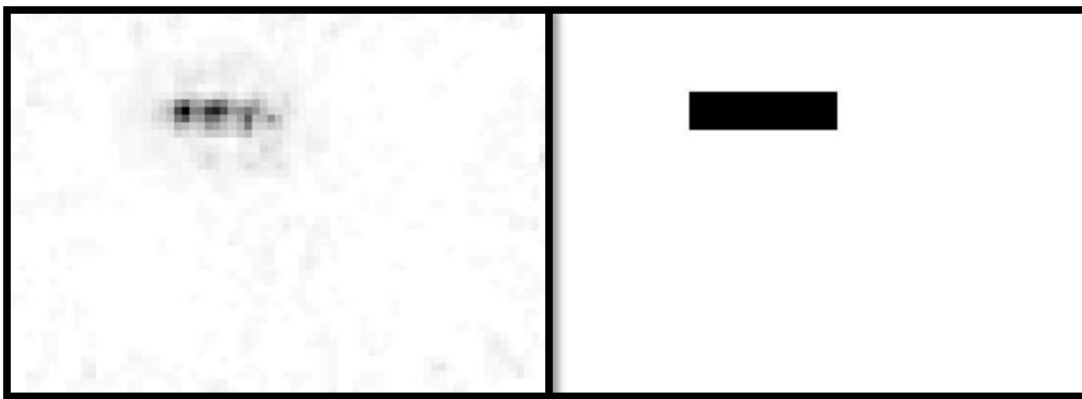


Figura 2.3: Imágenes generadas al azar.

2.2. Algoritmos de identificación de daños

Al momento de procesar las imágenes, se analizan distintas características, como por ejemplo la intensidad del pixel de estas para poder proceder a aplicar métodos de correlación

y así identificar las delaminaciones o fallas. A continuación, se mencionan dichos métodos. Cabe destacar que estos pueden ser usados de forma simultánea de tal forma de obtener resultados más óptimos.[6]

2.2.1. Promedio Factor de Correlación

Este método consiste en obtener un promedio de la intensidad de color de toda la imagen para luego multiplicarlo por un factor experimental (P) y obtener una cota superior de la intensidad de colores que se consideran como daño de las que no. Con esto se fija un rango y cualquier pixel o conjunto de ellos dentro de este rango se considera como una zona de posible delaminación. Cabe destacar que en este método al multiplicar el promedio por un factor determinado, estamos considerando una característica de cada una de las imágenes evitando casos particulares en donde puede ocurrir que los índices de daños se muestren bajo cierto umbral único.

Para obtener este valor P de forma óptima, se procede a calcular una Matriz de Confusión la cual consiste en obtener la proporción de Falsos Positivos y Falsos Negativos. Estos valores se describen a continuación:

Tabla 2.1: Matriz de Confusión.

		Situación Real	
		Dañada	No dañada
Predicción	Dañada	Verdadero Positivo (TP)	Falso Positivo (FP)
	No dañada	Falso Negativo (FN)	Verdadero Negativo (TN)

Luego de esta matriz se obtienen la proporción de los indicadores que se desea de la siguiente forma:

$$\text{Proporción de Falsos Positivos (FPR)} = \frac{FP}{FP + TN} \quad (2.1)$$

$$\text{Proporción de Falsos Negativos (FNR)} = \frac{FN}{FN + TP} \quad (2.2)$$

La FPR determina la proporción zonas no dañadas que son clasificadas incorrectamente como dañadas, y la FNR determina la proporción de zonas dañadas que son clasificadas incorrectamente como no dañadas. El objetivo de optimizar el valor de P es minimizar la FPR y FNR, para ello se toma una muestra de varias imágenes hasta obtener un valor tal que se adecue lo mayor posible a la mayor cantidad de imágenes muestreadas.

Para ello se realizar una iteración en donde el Factor P va variando, para cada Factor se define una matriz binario con 1 y 0 y se calculan los rates de Falsos Positivos y Negativos. Se realiza este paso para un número significativo de imágenes, se ponderan los resultados y se calcula el Factor promedio y el FNR dado un cierto FPR. En la Figura 2.4 se muestran la serie de pasos que se debe realizar para obtener un valor P y los valores outputs que se deben obtener.

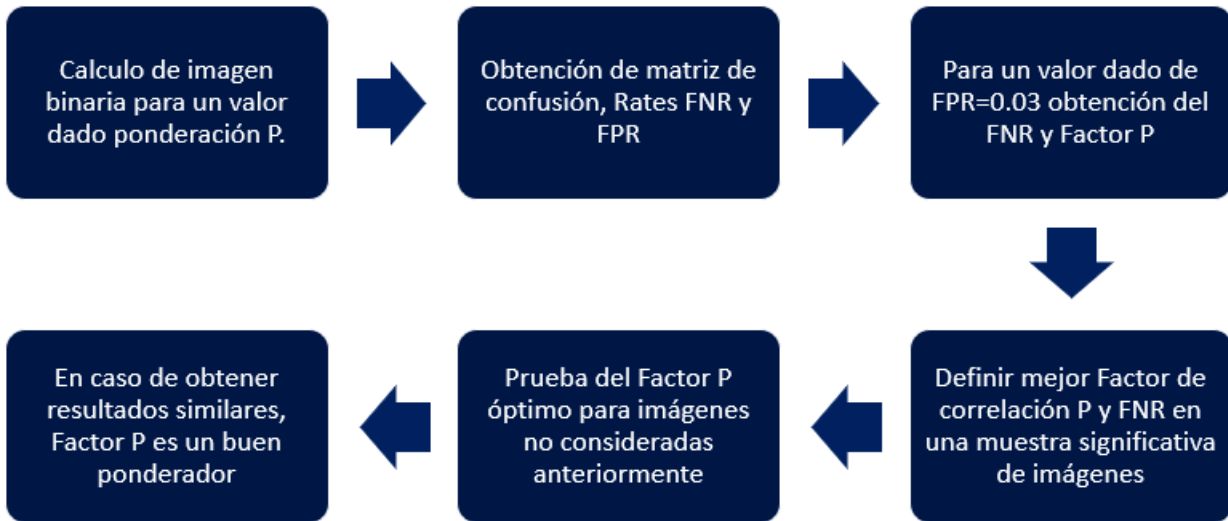


Figura 2.4: Metodología Promedio Factor Correlación

2.2.2. Filtros de menores tamaño

Este método es una mejora que se aplica a los algoritmo. Consiste en eliminar las manchas u objetos menores a cierto tamaño en la imagen. En este caso se define un tamaño de filtro de $2[mm]$. Con ello se cálcula la cantidad mínima de pixeles que deben tener las imágenes. Para fines prácticos y teniendo en cuenta una cantidad de pixeles conservadores, se obtiene dicho valor calculando el tamaño promedio de cada pixel en función del tamaño de la foto y el tamaño de la matriz que se genera. Se toma en cuenta un elemento circular y se define una cierta cantidad de pixeles mínimas que deben tener todos los objetos de las imágenes.

Debido a que es una mejora del algoritmo anterior se implementa este en dicho algoritmo. En la Figura 2.5 se muestra la serie de pasos que se debe realizar lograr obtener este Factor f para filtros.

2.2.3. Distribución T-Student Outliers

La idea de usar esta distribución es que la media y la desviación estándar de los índices de daño promedio en las regiones no dañadas de los paneles dañados son similares, por ello, suponiendo una distribución normal, los valores no típicos de la muestra serían considerados como índices de daños coincidentes con características estadísticamente significativas como el daño local [5].

El presente método adopta el test de Grubbs de construcción de mapeo de daño, el procedimiento consiste en [6]:

1. Primero, se obtiene el promedio (δ) y la desviación estándar (σ) de los índices de daños. Luego se obtiene el Residuo Máximo Normado (MNR) de la siguiente forma:

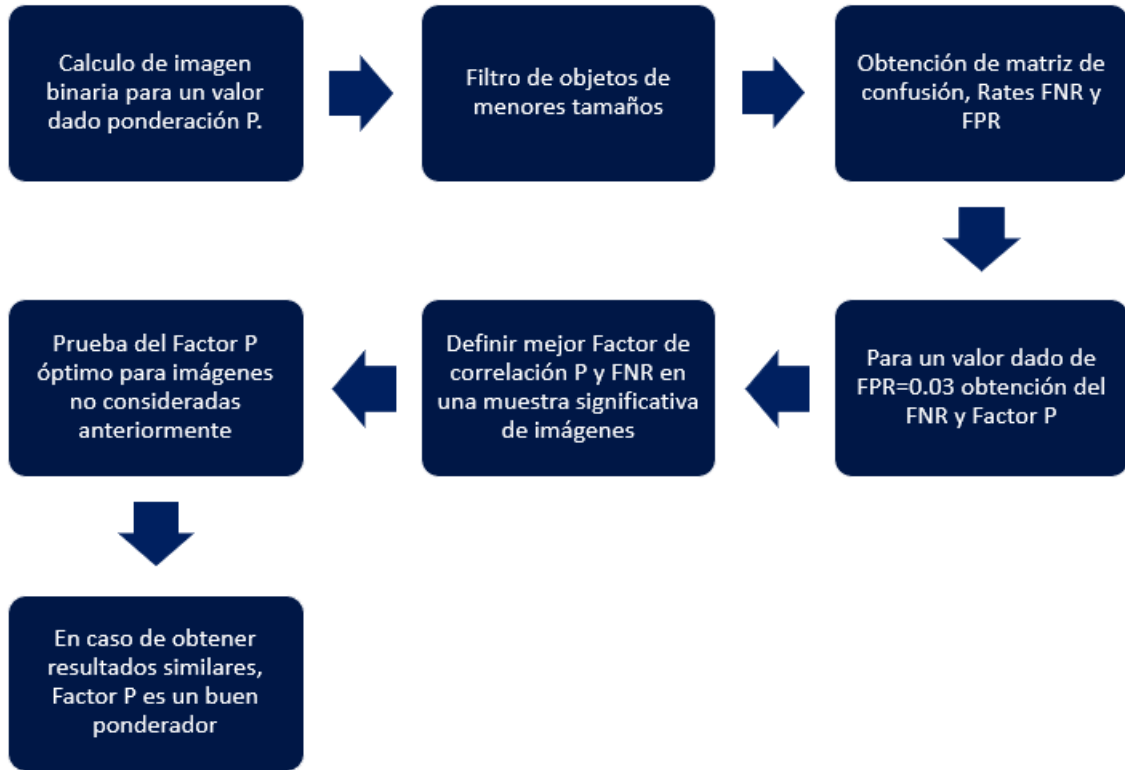


Figura 2.5: Metodología Promedio Factor Correlación con filtro

$$MNR = \max_k \left| \frac{\delta_k - \delta}{\sigma} \right| \quad (2.3)$$

En donde δ_k es el índice de daño en el punto k , donde $k=1, 2, \dots, n$ y n es el total de números de cuadrículas del indicador de daño.

2. Segundo se calcula un valor crítico (CV) el cual se compara con el MNR de la siguiente manera:

$$CV = \frac{t(n-1)}{\sqrt{n(n-2+t^2)}} \quad (2.4)$$

En donde t es el $1 - \frac{\alpha}{2n}$ quintil de la distribución T-Student con $n-2$ grados de libertad y con un nivel de significancia de 5% seleccionada.

3. Luego si es que MNR es más grande que CV , el punto asociado a MNR es declarado un outlier y es removido de la base de datos del indicador de daño. Se repite el paso 1 a 3 hasta que MNR se vuelva más pequeño que CV .
4. Finalmente, todos los puntos identificados como outlier son declarado como daños e incluidos en el mapa de daños. El resto de los puntos se indican como 1, es decir, zonas de no daño.

2.2.4. Proporción entre el área y el perímetro del contorno

El objetivo de esta correlación es definir una proporción entre área y perímetro, además de definir una relación en función del área y/o perímetro de la anomalía. [3]

Para ello primero se toma una muestra significativa de imágenes de daño y se realiza un gráfico entre áreas y perímetros. Con ello se define una Polifyt para obtener una función en donde se defina el comportamiento del área de las imágenes en función de su perímetro.

Luego de definir dicha imágenes se procesan imágenes en función del Factor óptimo de correlación P. Se calcula el área y perímetro de dicha imagen la cual se traduce en una cantidad de pixeles y se busca cambiar la forma de la imagen predefinida de tal forma que se adecue el perímetro a la función obtenida previamente manteniendo siempre la misma área.

En la Figura 2.6 se muestra la serie de pasos que se debe realizar. Cabe destacar que existe un paso previo al cambio de perímetro el cual es eliminar anomalías pequeñas las cuales no son consideradas en el estudio.

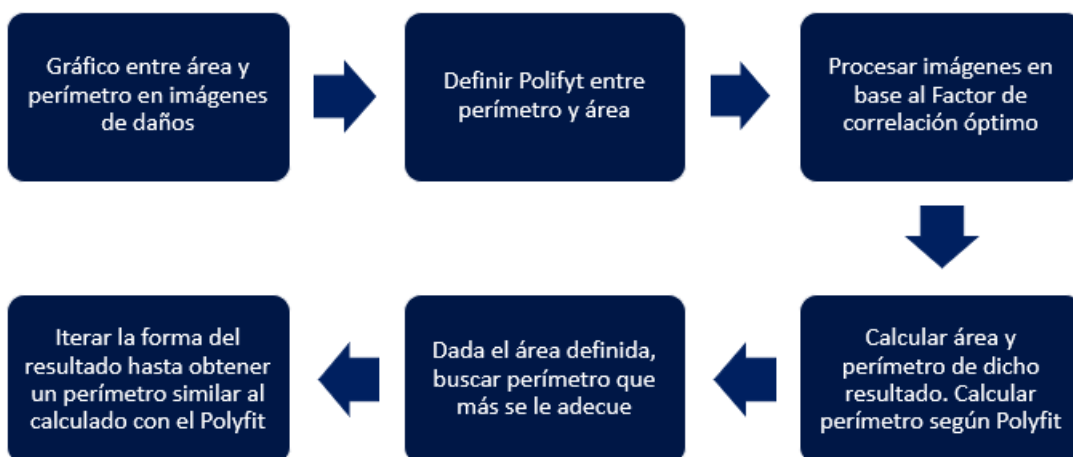


Figura 2.6: Metodología Proporción entre el área y el perímetro.

2.2.5. Proporción entre el área y el perímetro en base a elipses

Este método es similar al método nombrado anteriormente. La diferencia es que se definen elipses con el mismo segundo momento central normalizado y orientación que la imagen obtenida a través del método de ponderación. Con este método se busca obtener una mejor precisión al momento de buscar un Perímetro similar al definido por el Polifyt, dado que las imágenes obtenidas por diferentes métodos tienen un perímetro mucho más grande que un objeto cóncavo lo cual dificulta la convergencia de los perímetros.

El esquema en esta situación se muestra en la Figura 2.7 en donde sólo se adhiere un paso el cual consiste en el cálculo de cada elipse para cada objeto.

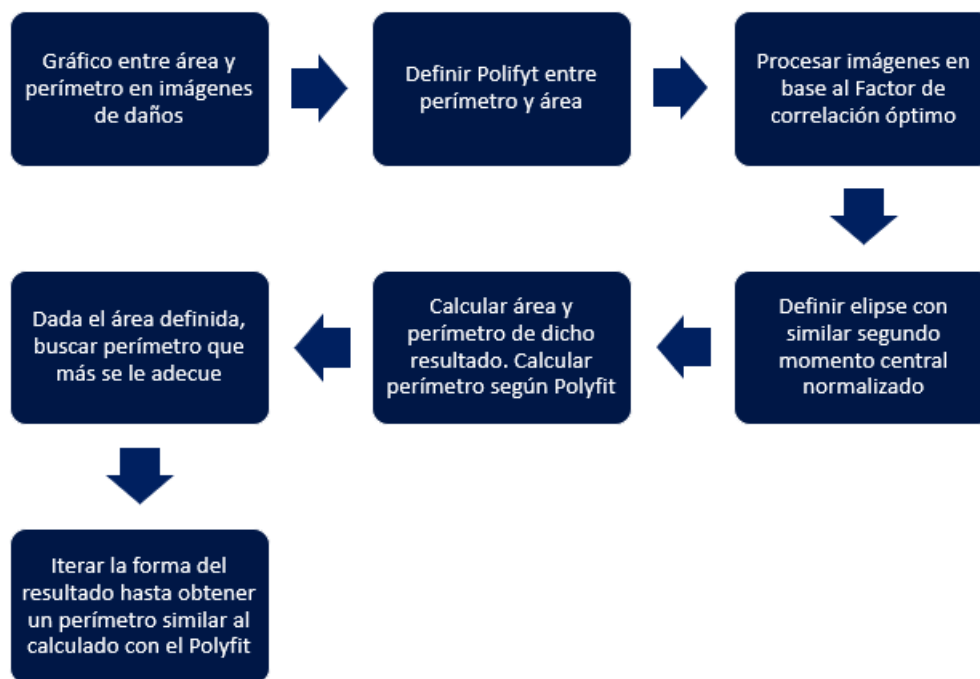


Figura 2.7: Metodología Proporción entre el área y el perímetro en base a elipses.

2.2.6. Variación de gradientes

Este método consiste en realizar una iteración hacia atrás para optimizar el cálculo del resultado previamente obtenido a través de cualquier método.[2]

Primero se procesa una imagen con índices de daño. Se compara el gradiente de cada pixel entre la imagen procesada y la imagen con índices de daño y luego en base a cierto umbral diferencia se redefine si dicho pixel corresponde a un daño real o no.

Es importante mencionar que hay lugares en donde el gradiente será muy grande debido que hay lugares en donde si no existe daño y si se considera y viceversa. En dichos casos se eliminan del muestreo de tal forma de reducir los Falsos negativos. En la Figura 2.8 se puede observar los pasos de esta metodología.

2.2.7. Filtro Área y Perímetro

Este método consiste un filtro adicional al método de Factor Ponderación con filtro, de tal forma de obtener un nuevo valor de ponderación P.[2]

El filtro adicional consiste que en base a la curva característica definida con ayuda del método, 2.2.4 Proporción entre el área y perímetro del contorno, se eliminan los objetos de las imágenes que se encuentren muy alejados de dicha curva.

Los pasos a seguir son similar al segundo algoritmo planteado solo que además de considerar el filtro por tamaño se realiza un filtro en base al área y perímetro de cada objeto, siendo

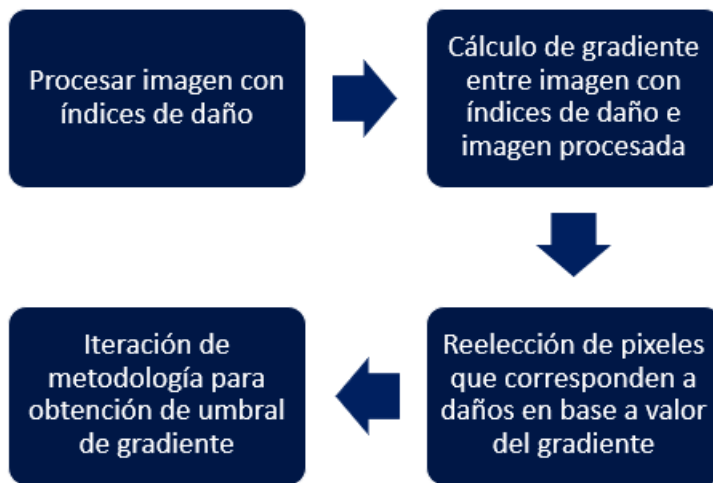


Figura 2.8: Metodología variación de gradiente.

eliminados los objetos cuya características estén muy alejadas de la curva. En la Figura 2.9 se puede observar los pasos de esta metodología.

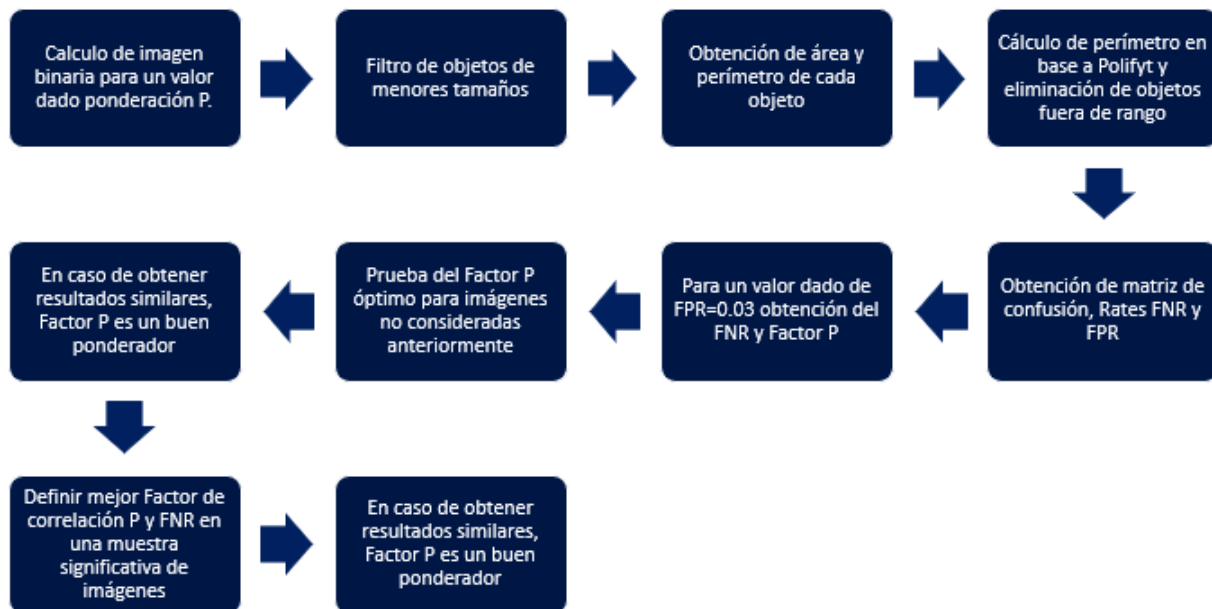


Figura 2.9: Metodología Filtro Área y Perímetro.

2.2.8. Imágenes experimentales

Para probar cada metodología de forma más real, se adquirieron imágenes tomadas de forma experimental y no simuladas. Dichas imágenes son obtenidas a través de la Profesora Guía del Trabajo de Título. En las Figuras 2.10, 2.11 y 2.12 se muestran los índices de daños de cada imagen junto con su daño real.

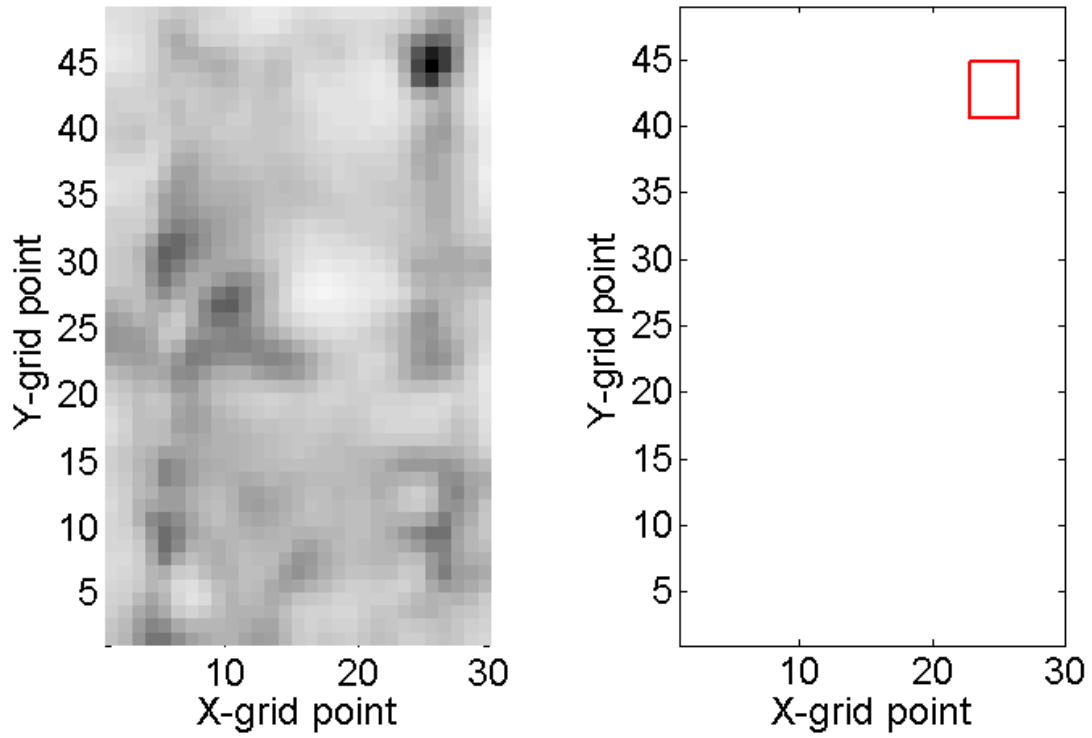


Figura 2.10: Imágenes experimentales, caso 1.

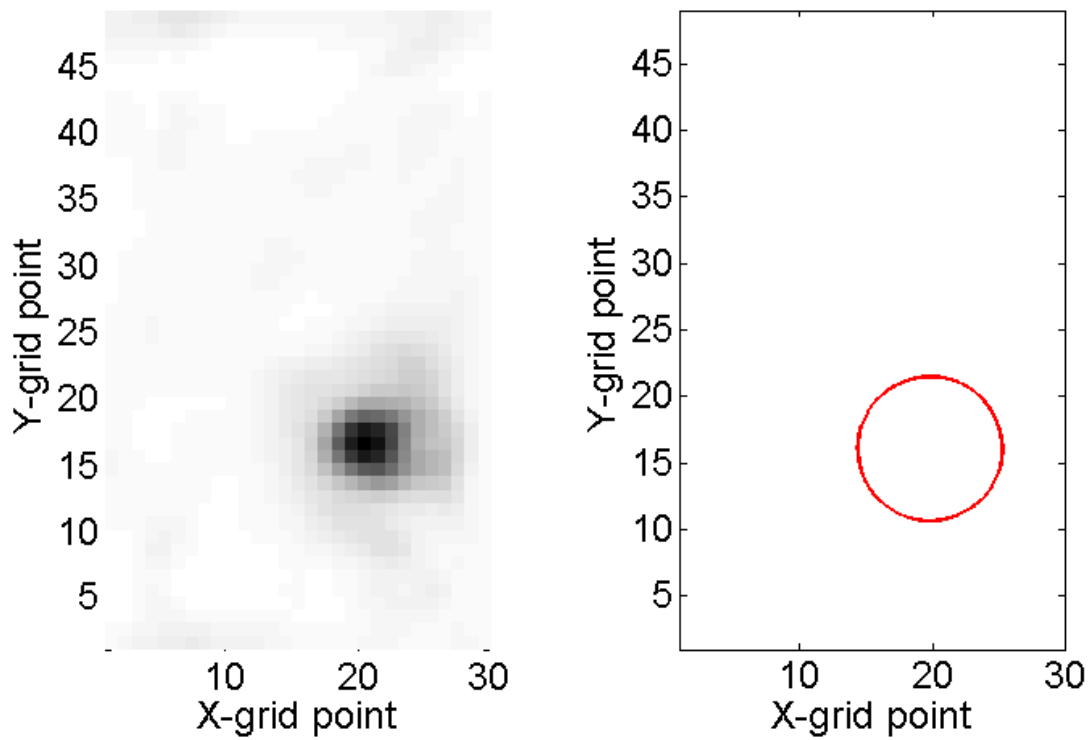


Figura 2.11: Imágenes experimentales, caso 2.

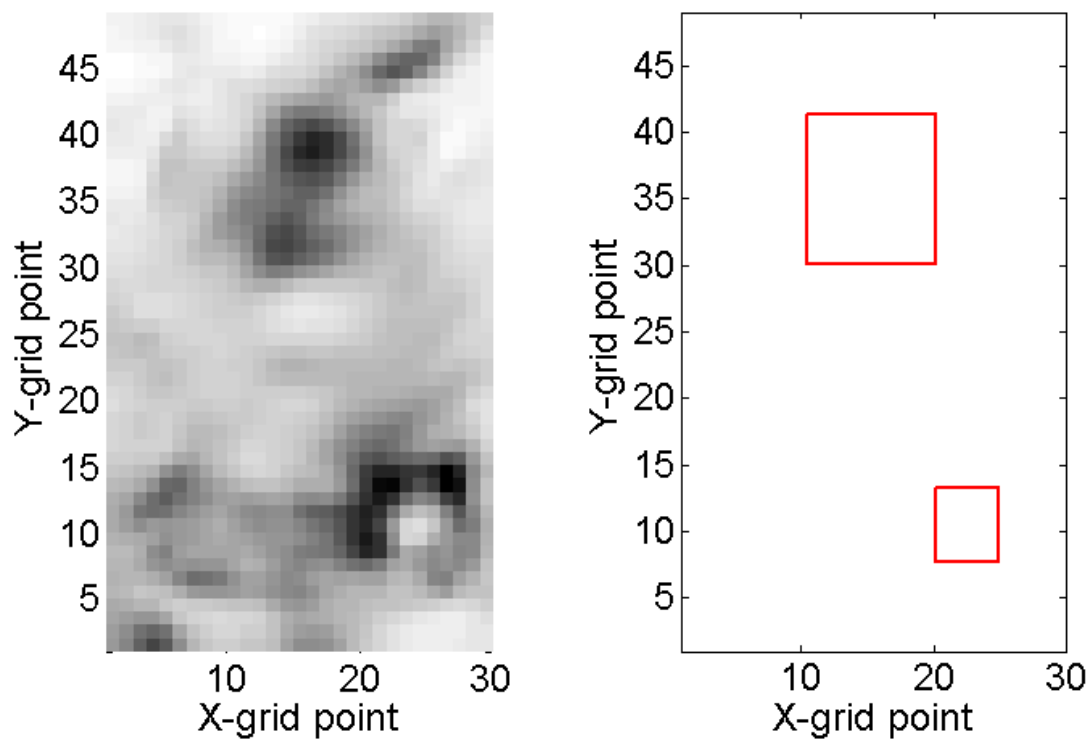


Figura 2.12: Imágenes experimentales, caso 3.

Capítulo 3

Metodología

Para el desarrollo de la memoria se propone la siguiente metodología.

1. Revisión de conceptos teóricos sobre las metodologías.

Al inicio del trabajo de memoria se realiza una revisión de la bibliografía recopilada para el desarrollo de esta, se busca información actualizada en artículos científicos y se genera una visión más específica de cómo proceder, es decir, se definen las metodologías a trabajar y cómo llegar a aplicarlas en forma correcta.

2. Análisis de metodologías de reconocimiento de imágenes aplicables al problema en estudio.

Se estudian y analizan las distintas metodologías para ver cómo aplicar cada una al problema en estudio y definir cómo implementar el algoritmo en MATLAB.

3. Implementación de las distintas metodologías en imágenes simuladas

Para un primer análisis se aplicarán métodos de procesamiento automático sobre imágenes obtenidas mediante una simulación en Elementos Finitos de la placa con daños creados al azar, con el fin de tener una gran variedad de muestras. Con ello se podrá estudiar y definir cuáles son las características más relevantes al momento de analizar una imagen a través de los algoritmos.

4. Análisis de sensibilidad de las metodologías.

Se analizan los parámetros y las funciones de los algoritmos de cada tipo de metodología y se seleccionan las características más óptimas de cada configuración para crear un algoritmo más completo y óptimo.

5. Aplicación y análisis de metodologías sobre imágenes experimentales.

Con una variedad de métodos de procesamiento ya establecidos, se estudiarán estos con imágenes tomadas de forma experimental y se analizarán los resultados obtenidos. De ser necesario, se realiza una iteración para mejorar el rendimiento de los métodos.

Cabe destacar que estas etapas se pueden realizar en forma paralela ya que al obtener una metodología eficiente se puede trabajar en otra para lograr una de mayor eficiencia u otra con resultados similares.

Capítulo 4

Resultados

A continuación, se muestran los resultados obtenidos en el Trabajo de Título.

4.1. Ponderación de Factor Correlación

En la Figura 4.1 se muestra un ejemplo del resultado de una imagen con índices de daños al aplicar este algoritmo. La imagen del centro es la imagen procesada con este algoritmo y la imagen de la derecha es la imagen que muestra el daño real.



Figura 4.1: a. Mapa de índices de daño, b. Daño detectado con Ponderación de Factores, c. Daño original.

El Factor de ponderación P obtenido es igual a 0,9437 utilizando 3000 imágenes simuladas. En la Figura 4.2 se puede observar la comparación de evolución de los rates de FNR y FPR en función de la variación del factor de ponderación. Cabe destacar que el resultado obtenido es un promedio y puede cambiar dependiendo de la imagen obtenida. Resulta interesante observar que aunque el Factor de ponderación P obtenido es único, el promedio de intensidad de los índices de daño es único para cada imagen, por ende, el resultado final que pondera ambos factores, es una propiedad interna de la imagen por lo cual aunque la imagen cambie en la intensidad de índices, siempre se tendrá una cota interna.

En la Tabla 4.7 se muestran los rates promedio de falsos negativos dado que se impone un rate FPR de 0.03, para los casos en donde el algoritmo se entrenó y se probó con imágenes

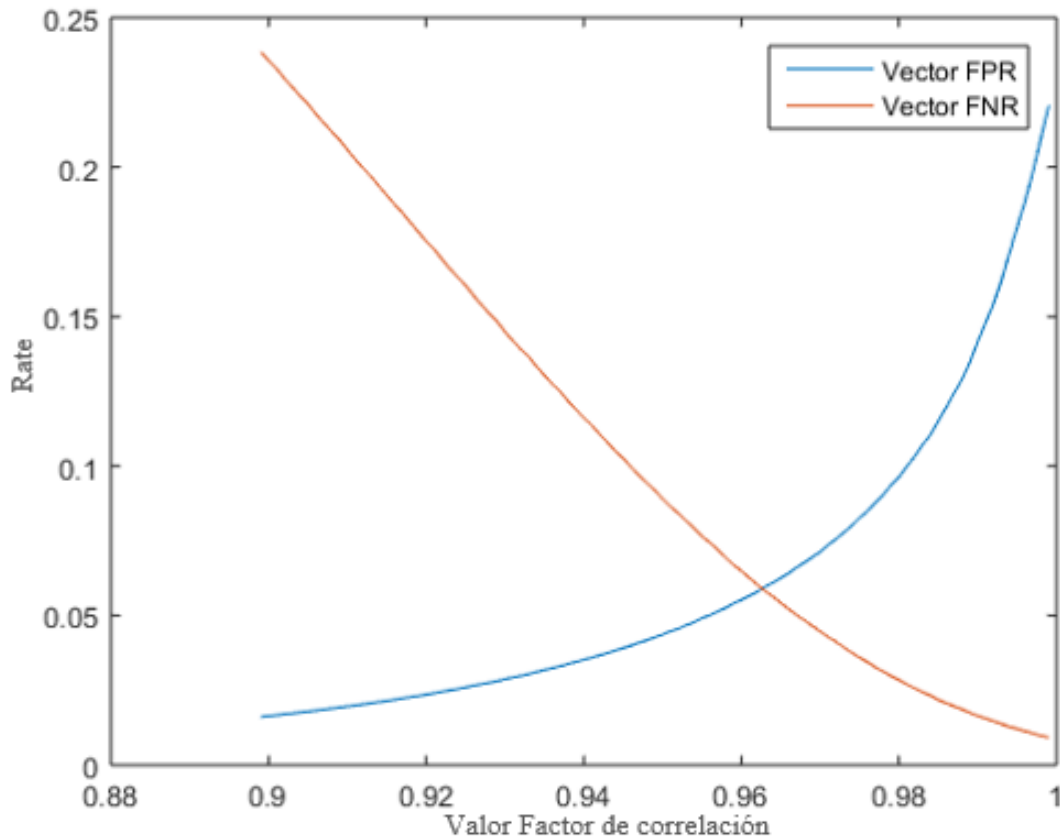


Figura 4.2: Evolución Rates de FPR y FNR en función del Factor de Correlación

creadas. Cabe destacar que la variación entre ambos FNR al ser un 12,65 % no resulta ser tan confiable, se lograba esperar un resultado de variación de un máximo 5%. Pero dada la certeza del rate FNR resulta ser bastante preciso comparado a otro métodos investigados.

Tabla 4.1: Rates obtenidos con Ponderación de Factor Correlación.

	Entrenamiento	Prueba imágenes generadas	Variación
FNR	0.0895	0.1002	12,65 %

En la Figura 4.3 se observa un histograma del rate FNR en donde se puede apreciar la variabilidad del algoritmo cuya desviación estándar es de 0,2276. Se aprecia que la cantidad de imágenes que tienen un FNR mayor al promedio son bastantes significativas por lo que con esto se comprueba que no es tan confiable esta metodología. También hay que recalcar que hay una gran cantidad de imágenes que tienen menor FNR, aunque sea mejor el resultado, dada la variabilidad de la metodología no es tan recomendable.

4.2. Filtro de menores tamaño

En la Figura 4.4 se muestra un ejemplo del resultado de una imagen con índices de daños al aplicar un filtro el cual elimina las "manchas" de menores tamaño junto con el algoritmo

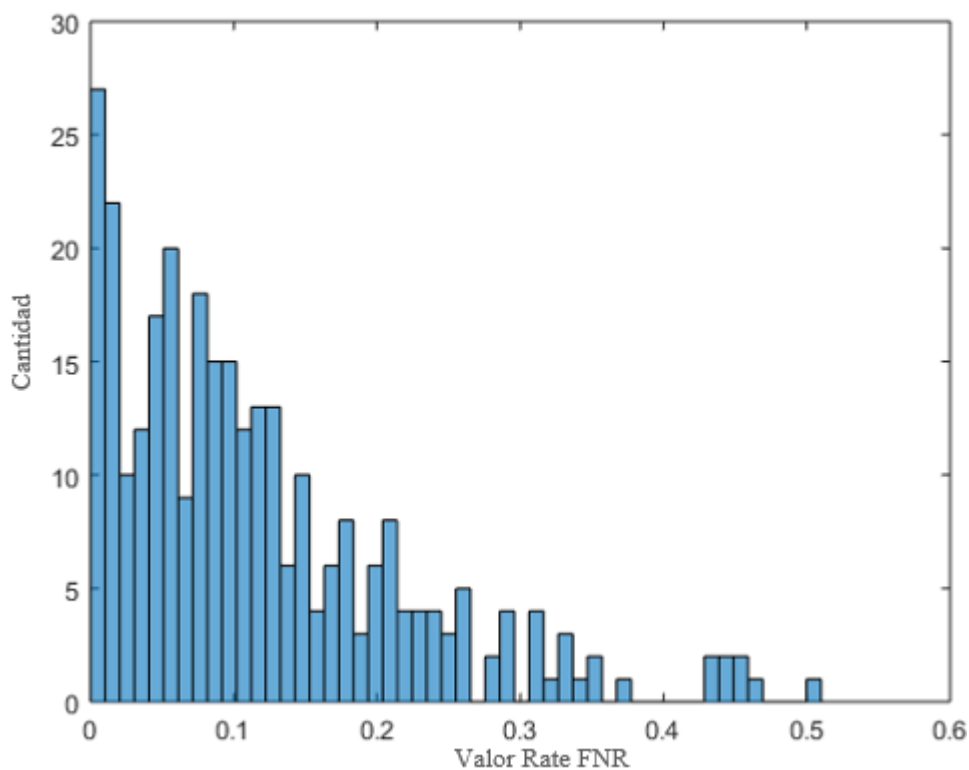


Figura 4.3: Histograma rate FNR con Ponderación de Factor Correlación.

de promedio de factor de correlación. La imagen del centro es la imagen procesada con este algoritmo y la imagen de la derecha es la imagen que muestra el daño real. A diferencia de la imagen obtenida con la primera metodología, esta no posee vacíos dentro del objeto de mayor envergadura, ni tanto existen "manchas"pequeñas dentro de la imagen.

Tomando en consideración que no se toman en cuenta manchas de tamaño menor a $2[cm]$ se calcula la cantidad máxima de píxeles que debe tener un objeto dentro de la imagen. Se obtiene un factor de 302 píxeles para aplicar el filtro. Con ello el Factor de ponderación P es igual a 0,9577.



Figura 4.4: a. Mapa de índices de daño, b. Daño detectado con Ponderación de Factores con filtro, c. Daño original.

En la Figura 4.5 se puede observar la comparación de evolución del los rates de FNR y FPR en función de la variación del factor de ponderación de los casos con y sin filtro. Se puede observar que en el caso con filtro los rates son menores lo que implica que el algoritmo

de Ponderación de Factor Correlación es más certero con esta mejora de filtros, ya que en el rate FPR=0,03 se logra una rate de FNR más alto que al no aplicar filtros.

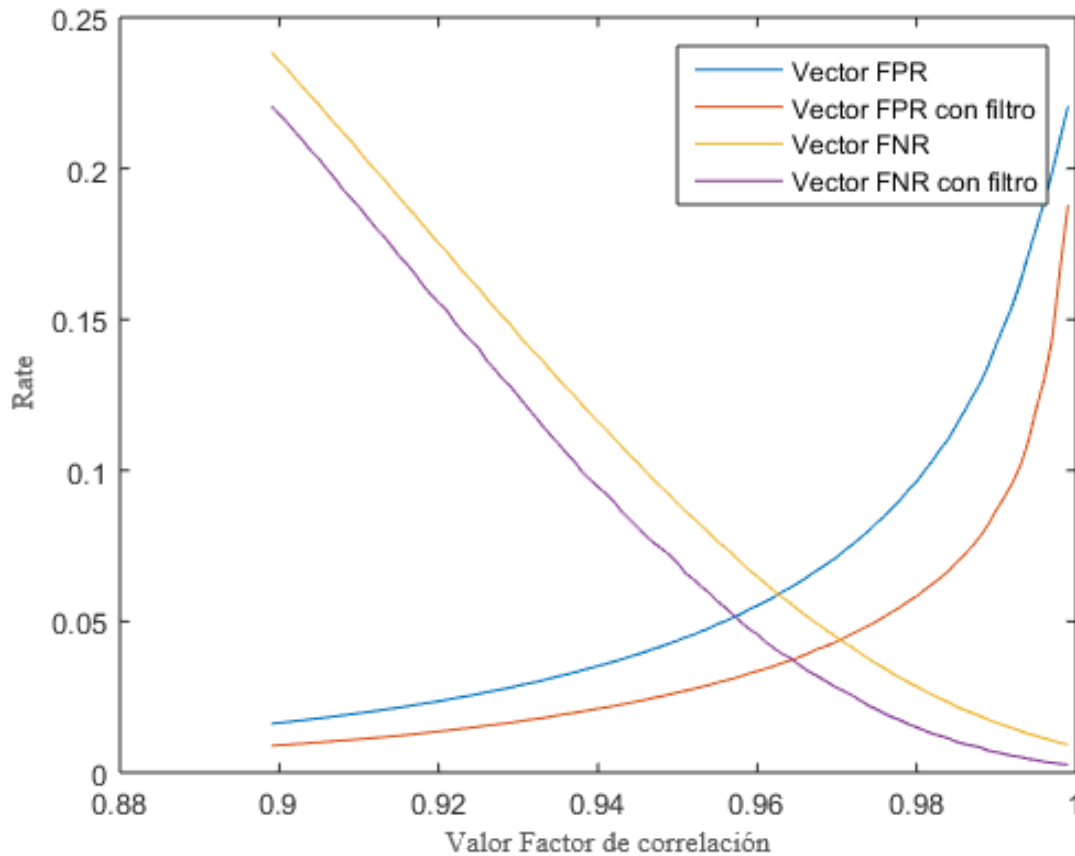


Figura 4.5: Evolución Rates de FPR y FNR en función del Factor de Correlación con y sin filtros.

En la Tabla 4.2 se muestran los rates promedio de falsos positivos y falsos negativos para los casos en donde el algoritmo se entrenó y se probó con imagenes creadas. Aquí se puede observar que se logró una variación de FNR menor al caso sin filtro y menor a un 5 %.

Tabla 4.2: Rates obtenidos con Ponderación de Factores con filtro.

	Entrenamiento	Prueba imágenes generadas	Variación
FNR	0,0431	0,0458	4,58 %

En la Figura 4.6 se observa un histograma del rate FNR en donde se puede apreciar la varibilidad del algoritmo cuya desviación estándar es de 0,1045. Al contrario del caso sin filtro se observa que la mayor cantidad de imagen esta concentrada en una sola zona. Se observa que existen imágenes que poseen un FNR muy lejando al promedio, pero de todas maneras resulta ser una mejora que le da mayor confiabilidad al algoritmo.

Dado el buen resultado obtenido a priori por aplicar filtros de manchas de menor tamaño, los próximos resultados se mostrarán en casos en donde se aplique o no se aplique dicho filtro.

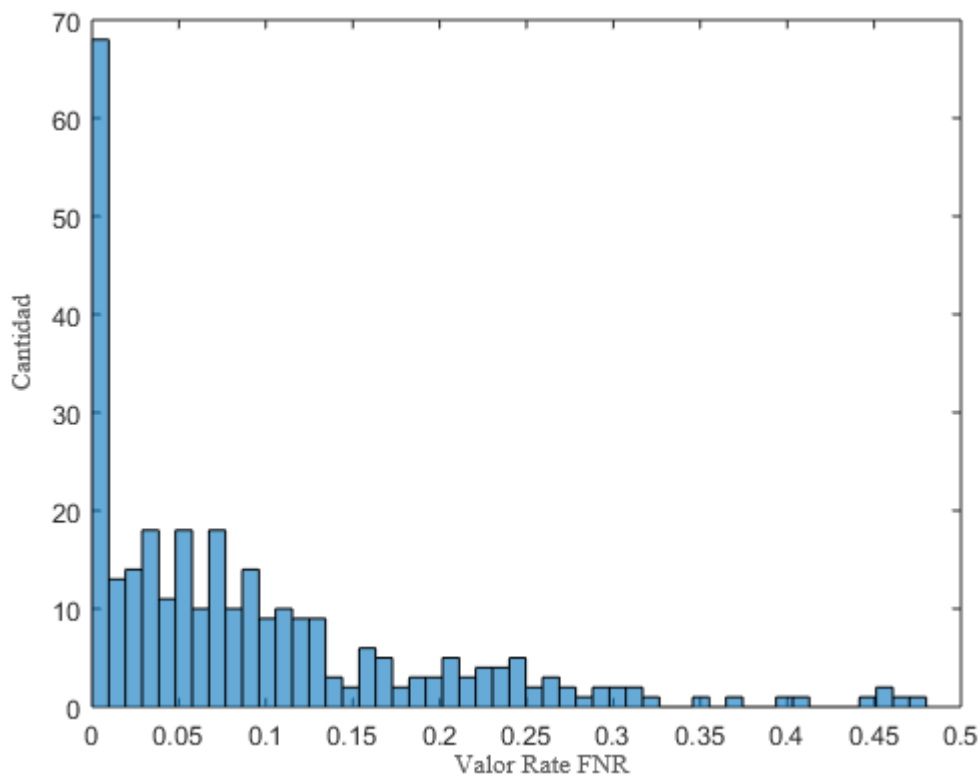


Figura 4.6: Histograma rate FNR con Ponderación de Factores con filtro.

4.3. Distribución T-Student Outliers

En la Figura 4.7 se muestra un ejemplo del resultado de una imagen con índices de daños al aplicar este algoritmo. La imagen del centro es la imagen procesada con este algoritmo y la imagen de la derecha es la imagen que muestra el daño real. El factor de significancia usado es de 5%.

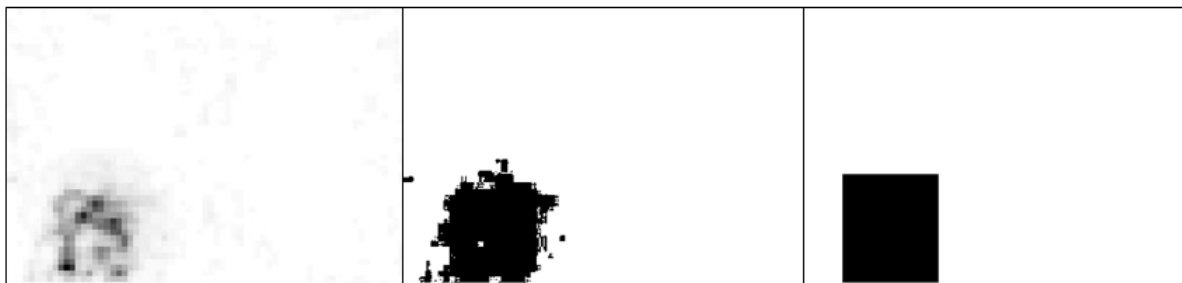


Figura 4.7: a. Mapa de índices de daño, b. Daño detectado con Método T-Student, c. Daño original.

En la Tabla 4.3 se muestran los rates promedio de falsos positivos y falsos negativos para los casos en donde el algoritmo se probó con imagenes creadas. Se puede observar que los rates de FNR son mucho más elevados que con el método de Ponderación con Factor Correlación.

En la Figura 4.8 se observa un histograma del rate FNR en donde se puede apreciar la

Tabla 4.3: Rates obtenidos con Método T-Student.

FNR	FNR con filtro	FPR	FPR con filtro
0.2048	0.2006	0.02	0.0163

variabilidad del algoritmo cuya desviación estándar es de 0,1937. Se logra apreciar una mayor variabilidad que en el método de Ponderación de Factor Correlación.

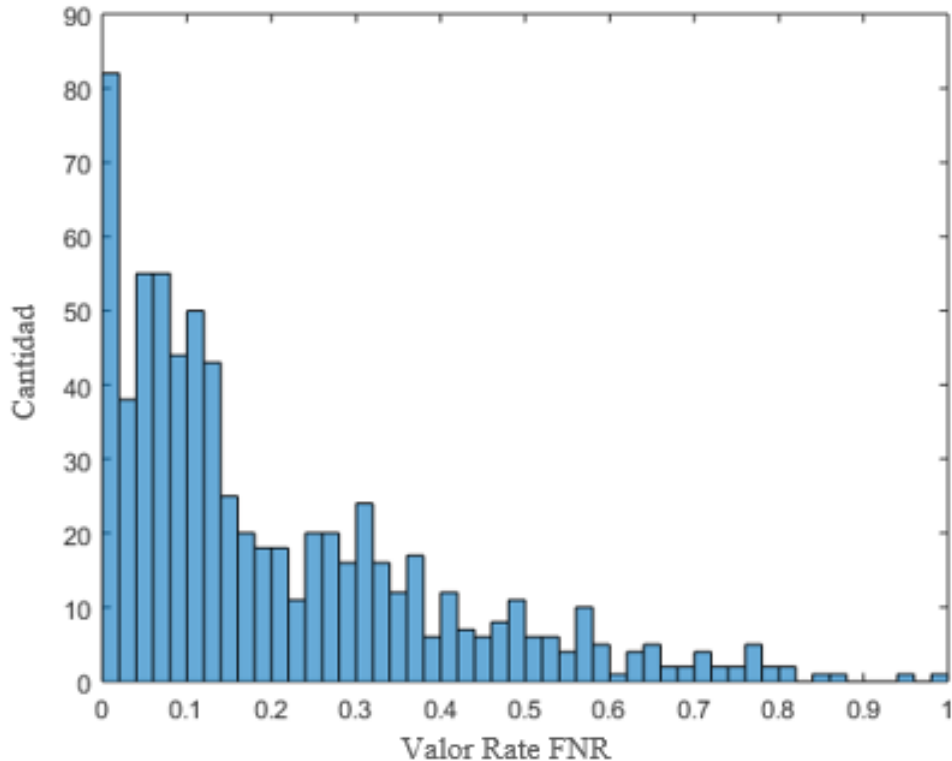


Figura 4.8: Histograma rate FNR con Método T-Student sin filtro.

En la Figura 4.9 se observa un histograma del rate FNR en donde se puede apreciar la variabilidad del algoritmo cuya desviación estándar es de 0,2326. En este caso con filtro se puede apreciar que la variabilidad disminuye de gran manera, pero el rate FNR no logra disminuir de gran manera.

4.4. Proporción entre el área y el perímetro

En la Figura 4.10 se muestra un ejemplo del contorno de la imagen obtenida junto con la solución del daño que se quiere obtener. La imagen del centro es la imagen procesada con este algoritmo y la imagen de la derecha es la imagen que muestra el daño real.

En la Figura 4.11 se muestra el Jackknife entre las área y perímetro de una muestra significativa de imágenes y el polinomio característico de los datos graficados. Con ellos se obtuvo que la función más certera es un polinomio de grado 2 de área en función del perímetro: $0,0642x^2 + 2,03x + 181$.

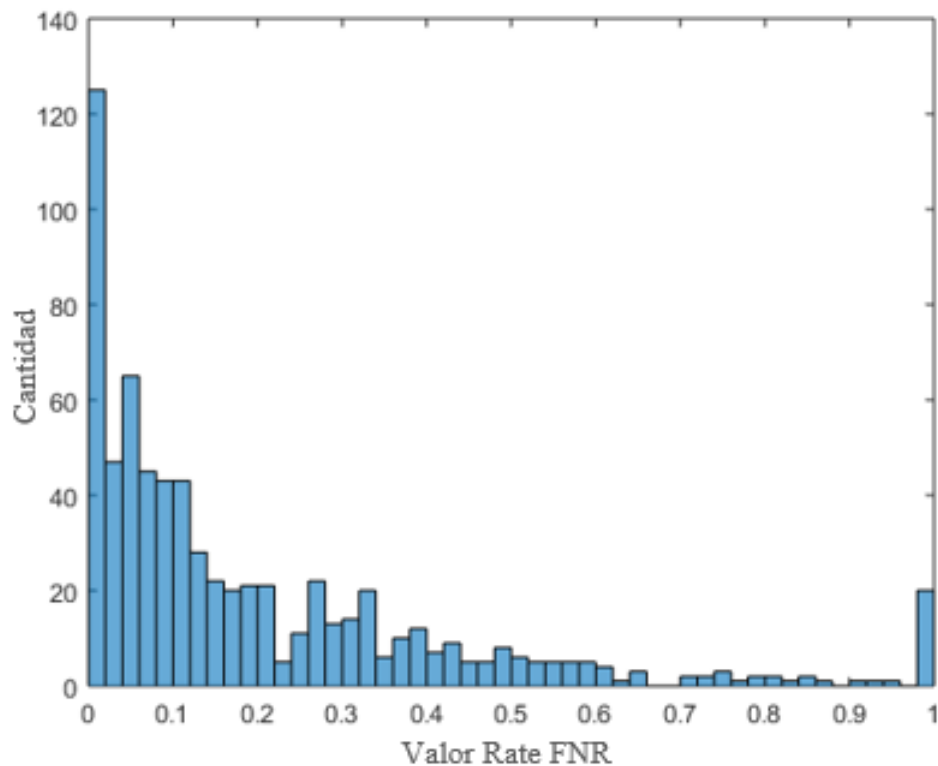


Figura 4.9: Histograma rate FNR con Método T-Student con filtro.

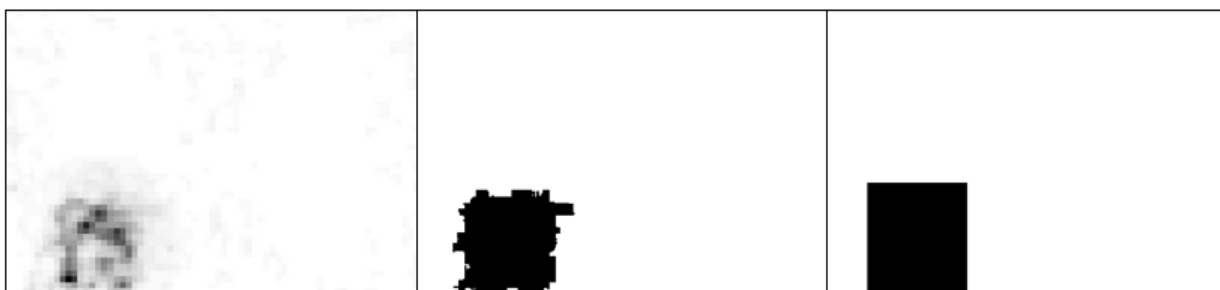


Figura 4.10: a. Mapa de índices de daño, b. Daño detectado en base a Proporción entre Área y Perímetro, c. Daño original.

En la Tabla 4.4 se muestran los rates promedio de falsos positivos y falsos negativos para los casos en donde el algoritmo se probó con imagenes creadas. Dado que la base de este algoritmo es reordenar la cantidad de pixeles obtenidos en el caso anterior de tal forma que se obtenga un menor rate del FNR, se aprecia que el rate FPR es muy similar al impuesto en el primer proceso (FPR=0.03). No así el rate FNR, ya que este aumento de un valor de 0,0431 a 0,053, por lo que pierde certeza.

Tabla 4.4: Rates obtenidos en base a Proporción entre Área y Perímetro.

FNR	FPR
0,053	0,0303

En la Figura 4.12 y 4.13 se observa un histograma del rate FNR y FPR en donde se

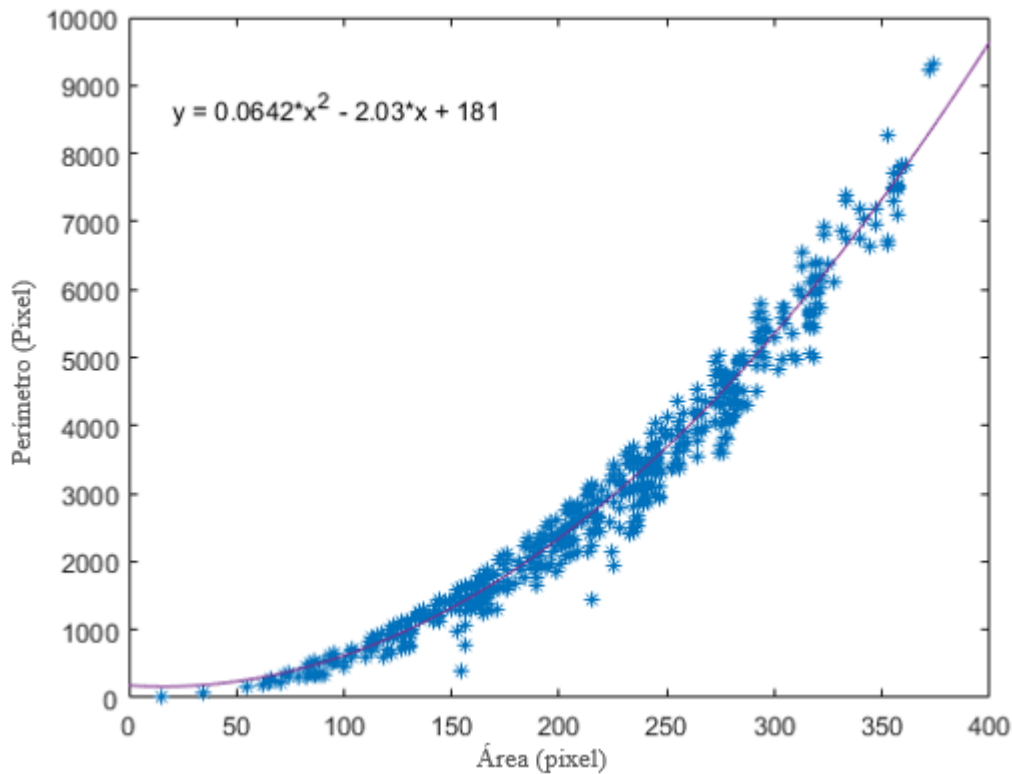


Figura 4.11: Función de relación entre área y perímetro.

Se puede apreciar la variabilidad de estos parámetros cuyas desviaciones estándar son de 0,0178 y 0,0637. Se aprecia que el rate FPR tiene una gran variabilidad dada que la situación inicial era nula ya que había impuesto un valor de 0,03 (Método de Ponderación con Factor Correlación). No así la variabilidad del rate FNR, ya que comparada con la Figura 4.6 en la cual se aprecia que la variabilidad es mayor.

4.5. Proporción área y perímetro modelada con elipses

En la Figura 4.14 se muestra un ejemplo del contorno de la imagen obtenida junto con la solución del daño que se quiere obtener. La imagen del centro es la imagen procesada con este algoritmo y la imagen de la derecha es la imagen que muestra el daño real.

En la Figura 4.15 se muestra el Jackknife entre las área y perímetro de una muestra significativa de imágenes y el polinomio característicos de los datos graficados. Con ellos se obtuvo que la función más certera es un polinomio de grado 3: $0,000249x^3 - 0,0439x^2 + 16,5x - 661$

En la Tabla 4.6 se muestran los rates promedio de falsos positivos y falsos negativos para los casos en donde el algoritmo se probó con imágenes creadas. Se observa que se logra obtener un FNR menor con el mismo FPR que el caso en que no se modela con elipses.

En la Figura 4.16 y 4.17 se observa un histograma del rate FNR y FPR en donde se

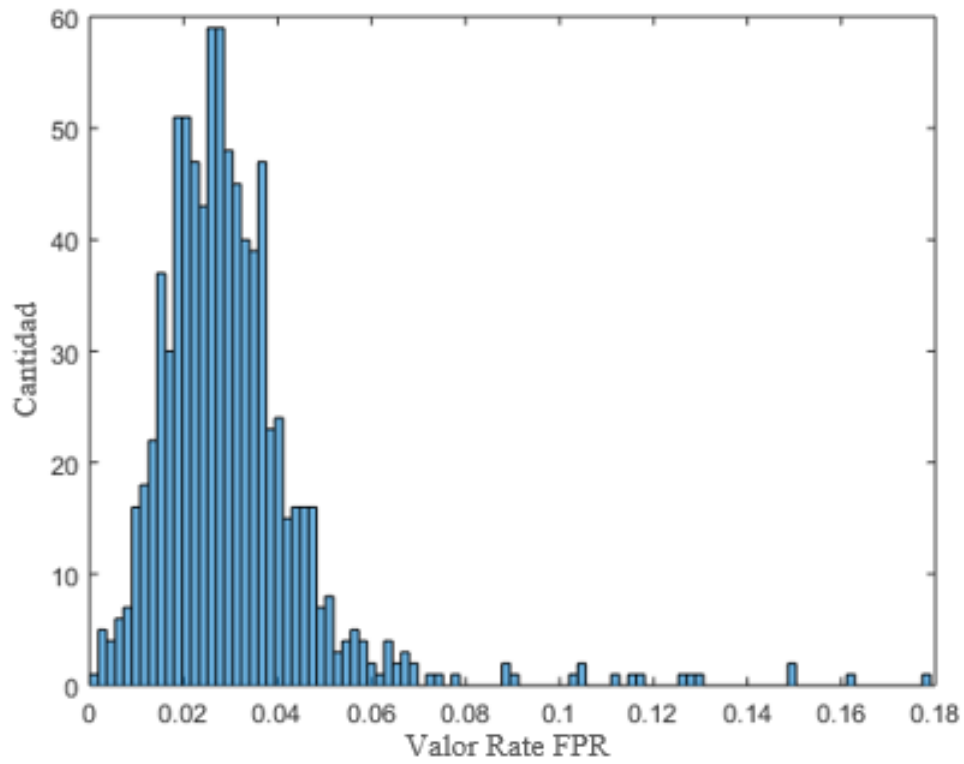


Figura 4.12: Histograma rate FPR en base a Proporción entre Área y Perímetro.

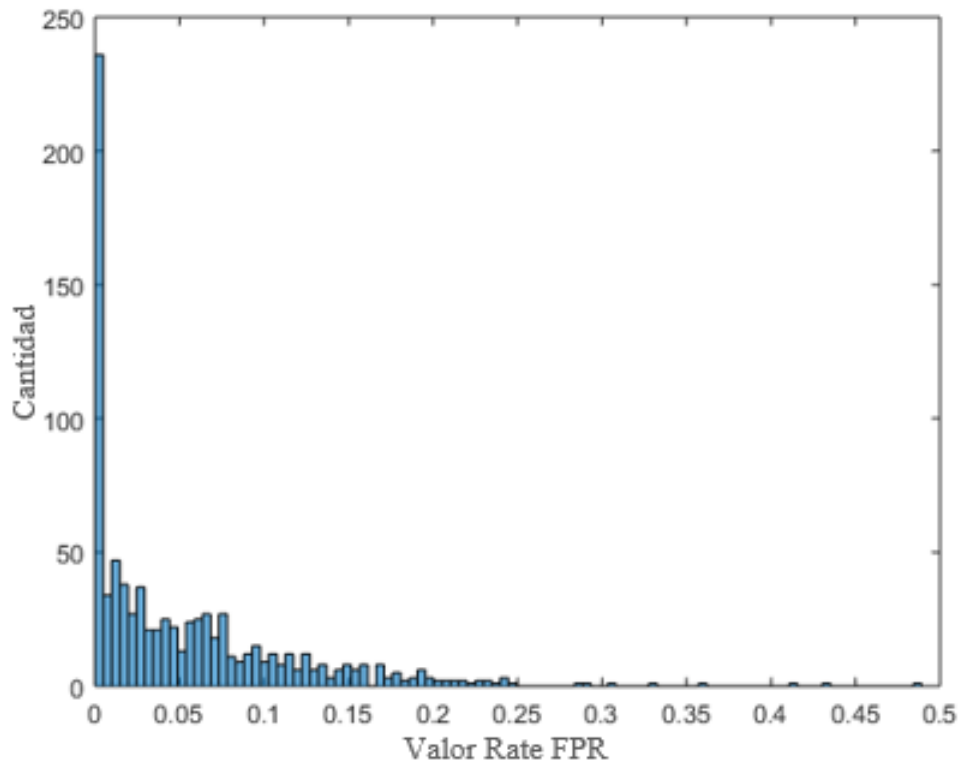


Figura 4.13: Histograma rate FNR en base a Proporción entre Área y Perímetro.

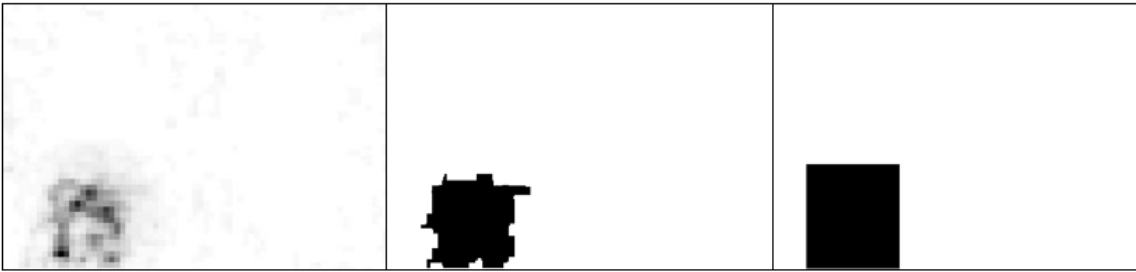


Figura 4.14: a. Mapa de índices de daño, b. Daño detectado en base a Proporción entre Área y Perímetro modelada con elipses, c. Daño original.

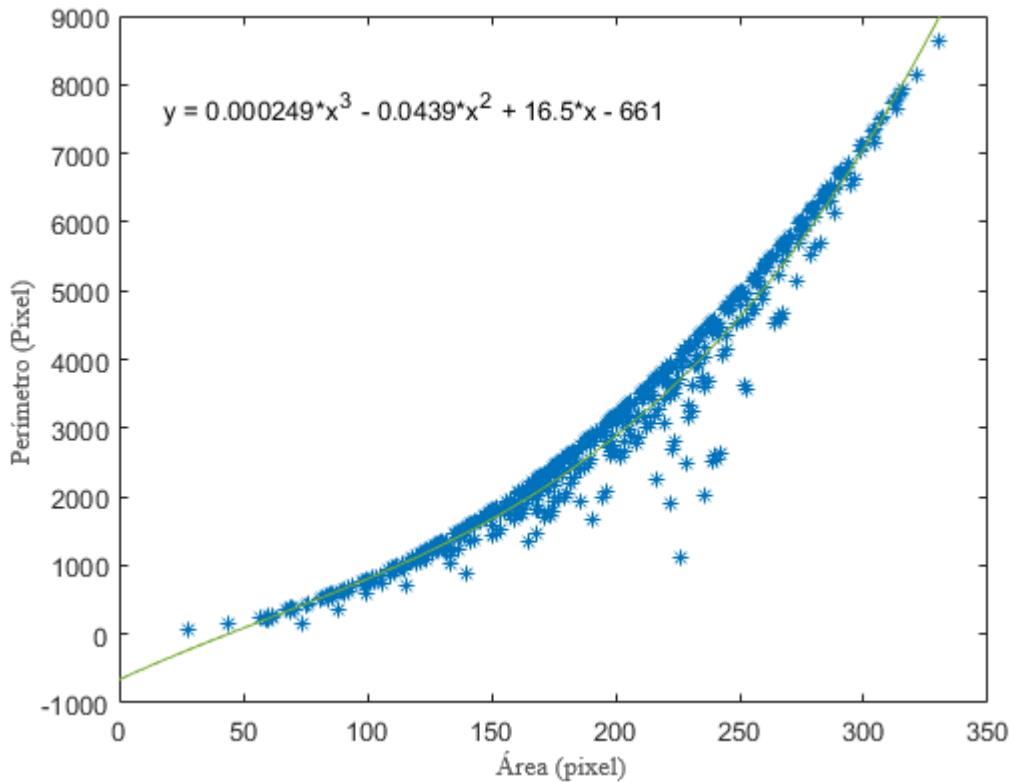


Figura 4.15: Función de relación entre área y perímetro de elipses.

Tabla 4.5: Rates obtenidos en base a Proporción entre Área y Perímetro modelada con elipses.

FNR	FPR
0,0519	0,0313

puede apreciar la variabilidad de estos parámetros cuyas desviaciones estándar son de 0,0211 y 0,0623. Se observa que los gráficos son muy similares, lo cual indica que no se logra mejorar de forma significativa la variabilidad entre el método modelado con elipses y sin elipses.

En la Figura 4.18 se puede apreciar un ejemplo en donde se muestra la elipse con el mismo segundo momento central normalizado. Cabe destacar que en base a esa elipse definida se hicieron los cálculos obtenidos anteriormente. También se puede apreciar el centro de masa del objeto el cual es el mismo tanto para el objeto demorado y para la elipse con el mismo

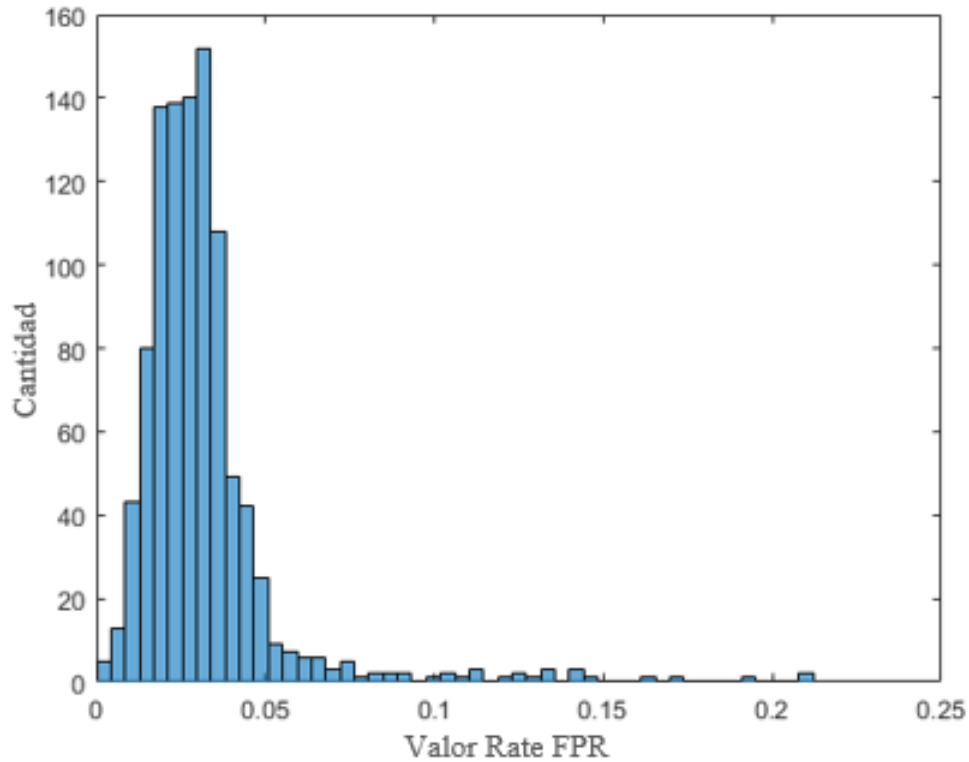


Figura 4.16: Histograma rate FPR en base a Proporción entre Área y Perímetro modelada con elipses.

momento.

4.6. Variación de gradientes

En la Figura 4.19 se muestra un ejemplo del contorno de la imagen obtenida junto con la solución del daño que se quiere obtener. La imagen del centro es la imagen procesada con este algoritmo y la imagen de la derecha es la imagen que muestra el daño real.

En la Tabla 4.6 se muestran los rates promedio de falsos positivos y falsos negativos para los casos en donde el algoritmo se probó con imagenes creadas. Se observa que se logra obtener un FNR menor con el mismo FPR que el caso en que no se modela con elipses.

Tabla 4.6: Rates obtenidos en base a Variación de gradientes.

FNR	FPR
0,0946	0,031

El Factor de ponderación PG para obtener el valor del umbral de diferencia valor mínimo de gradiente es igual a 0,9156. En la Figura 4.20 se puede observar la comparación de evolución del los rates de FNR y FPR en función de la variación del factor de ponderación. Cabe destacar que el resultado obtenido es un promedio y puede cambiar dependiendo de la imagen obtenido.

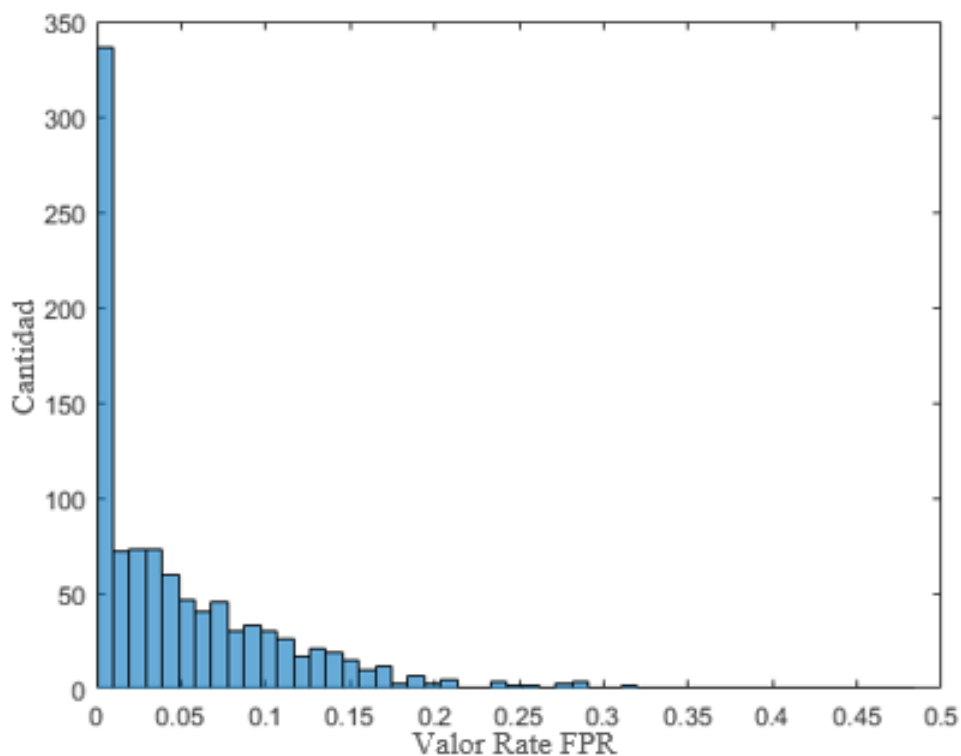


Figura 4.17: Histograma rate FNR en base a Proporción entre Área y Perímetro modelada con elipses.

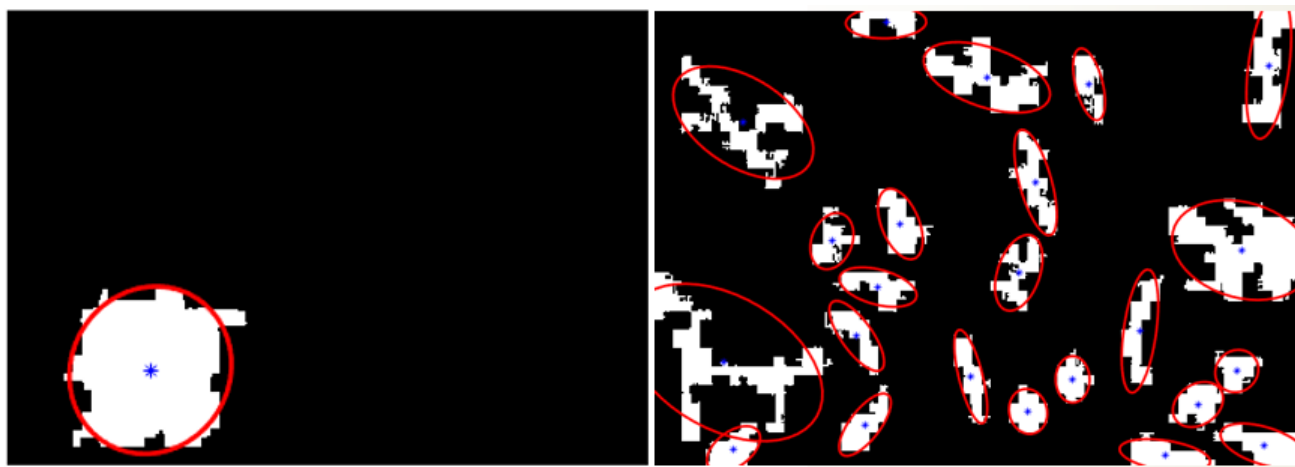


Figura 4.18: Histograma rate FPR.

4.7. Filtro Área y Perímetro

En la Figura 4.27 se muestra un ejemplo del resultado de una imagen con índices de daños al aplicar este algoritmo. La imagen del centro es la imagen procesada con este algoritmo y la imagen de la derecha es la imagen que muestra el daño real.

El Factor de ponderación P obtenido es igual a 0,9856 utilizando 3000 imágenes simu-



Figura 4.19: a. Mapa de índices de daño, b. Daño detectado en base a Variación de Gradientes, c. Daño original.

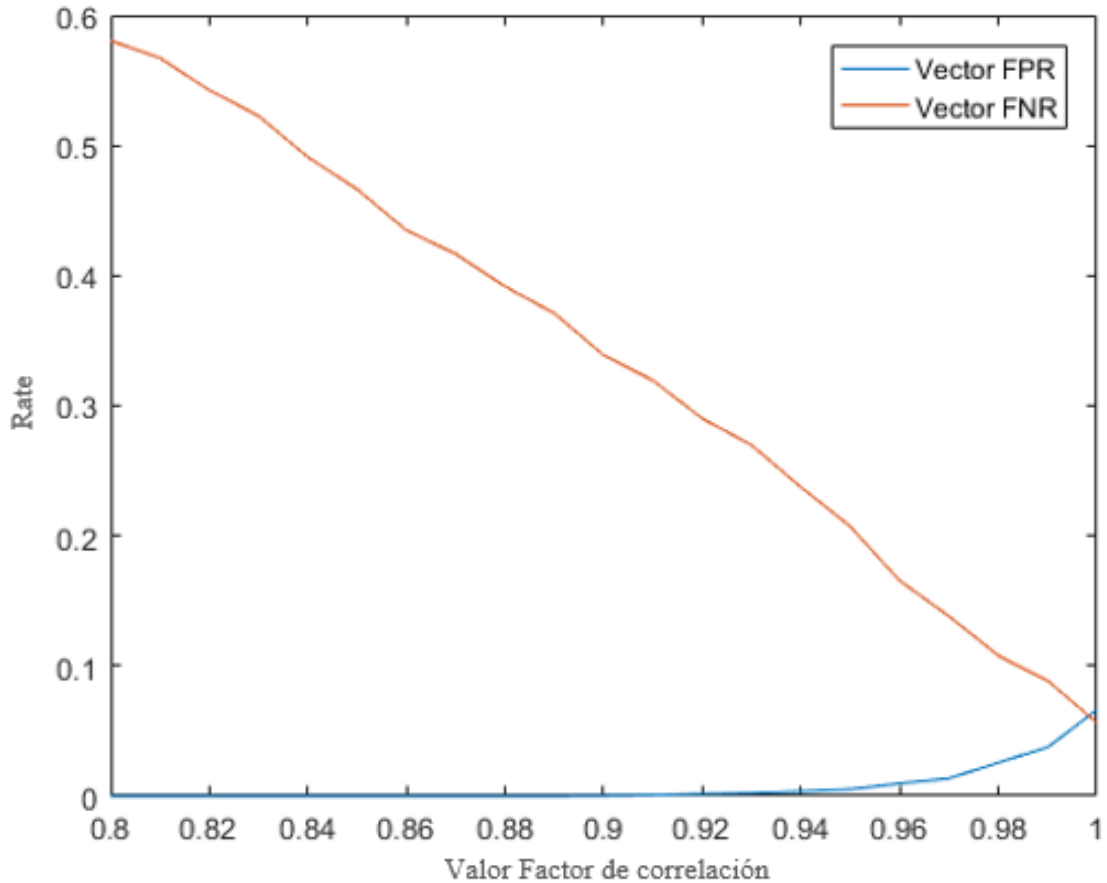


Figura 4.20: Evolución Rates de FPR y FNR en base a Variación de gradientes.

ladas. Como se dijo anteriormente el resultado obtenido es un promedio y puede cambiar dependiendo de la imagen obtenida, y además resultado final es una ponderación de dos factores, el cual uno de ellos, es una propiedad interna de la imagen por lo cual aunque la imagen cambie en la intensidad de índices, siempre se tendrá una cota interna.

En la Tabla 4.7 se muestran los rates promedio de falsos negativos dado que se impone un rate FPR de 0.03, para los casos en donde el algoritmo se entrenó y se probó con imágenes creadas. Cabe destacar que la variación entre ambos FNR al ser un 4,27% resulta ser un poco menor a la variación con solo el filtro de tamaño.

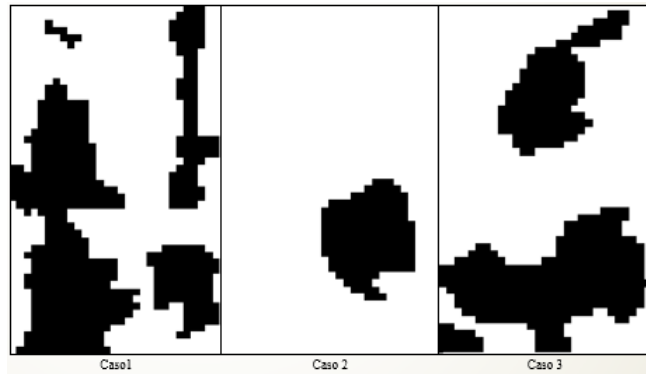


Figura 4.21: a. Mapa de índices de daño, b. Daño detectado con Ponderación de Factores con filtros de tamaño y área-perímetro, c. Daño original.

Tabla 4.7: Rates obtenidos con Ponderación de Factor Correlación con filtros de tamaño y área-perímetro.

	Entrenamiento	Prueba imágenes generadas	Variación
FNR	0.1039	0,1083	4,27 %

En la Figura 4.22 se observa un histograma del rate FNR en donde se puede apreciar la variabilidad del algoritmo cuya desviación estándar es de 0,0669.

En la Tabla 4.8 se puede apreciar un resumen de los valores de FNR y FPR promedios obtenidos para en imágenes experimentales procesadas por cada algoritmo.

Tabla 4.8: Valores de FNR y FPR promedios de cada algoritmo en imágenes simuladas.

Algoritmo	FNR	FPR
1. Promedio Factor Correlación	0,100	0,030
2. Promedio Factor Correlación con filtro	0,046	0,030
3. Distribución T-Student	0,205	0,020
4. Distribución T-Student con filtro	0,201	0,016
5. Proporción Área-Perímetro	0,053	0,030
6. Proporción Área-Perímetro en base a elipses	0,052	0,031
7. Variación gradientes	0,095	0,031
8. Promedio Factor Correlación con dos filtros	0,104	0,108

4.8. Algoritmos en imágenes reales

A continuación se muestran las imágenes tomadas en forma experimental procesadas por los diferentes algoritmos que se han mencionado a lo largo de este trabajo. Cabe destacar que no se obtuvieron resultados para la imagen caso 3 con el Método T-student.

Se puede apreciar que no existen grandes diferencias entre imágenes procesadas en base a proporción y área con respecto a Ponderación Factor Correlación.

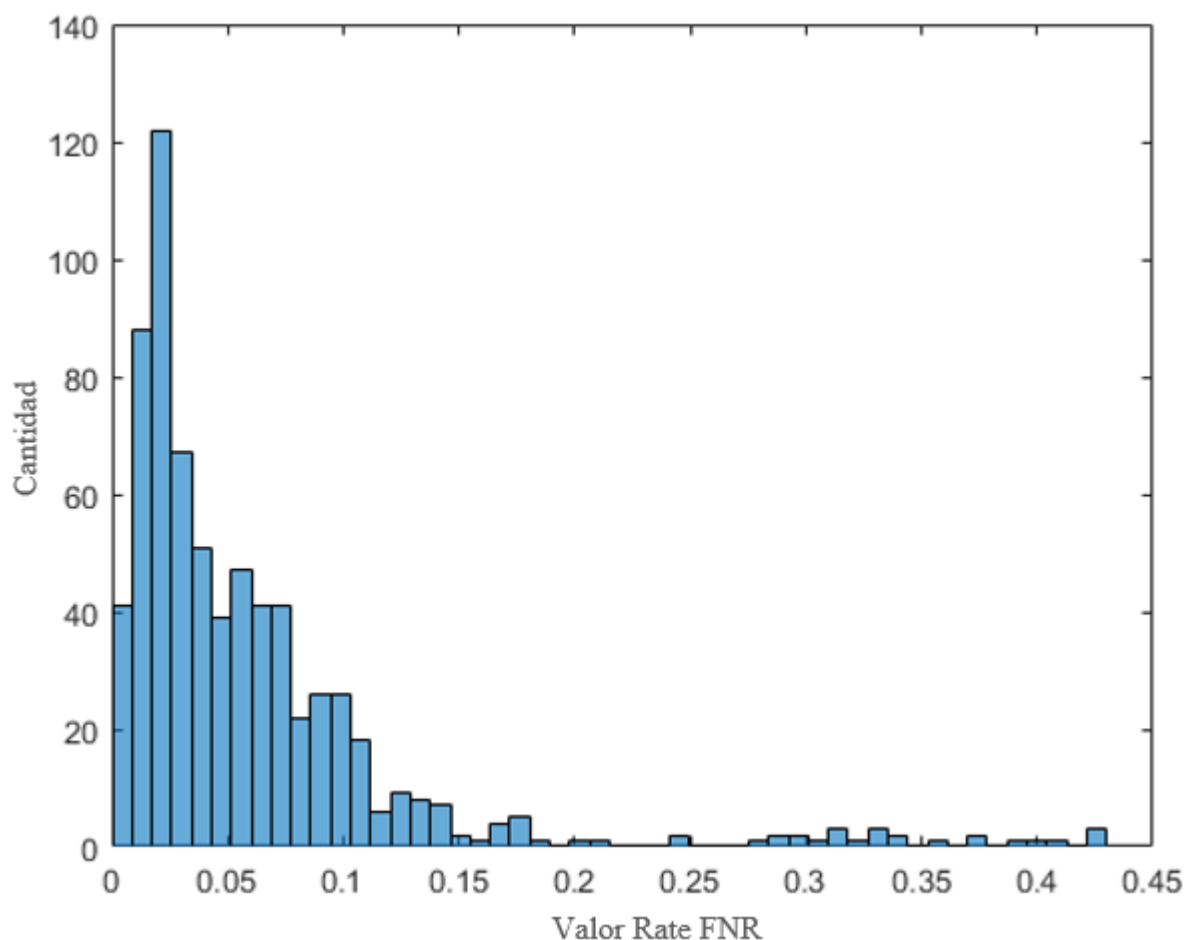


Figura 4.22: Histograma rate FNR con Ponderación de Factor Correlación con filtros de tamaño y área-perímetro.

En la Tabla 4.9 se puede apreciar un resumen de los valores de FNR y FPR obtenidos para cada caso de imágenes experimentales procesadas por cada algoritmo.

Tabla 4.9: Valores de FNR y FPR para cada caso por cada algoritmo en imágenes experimentales.

	Caso 1		Caso 2		Caso 3	
	FNR	FPR	FNR	FPR	FNR	FPR
1. Promedio Factor Correlación	0,461	0,216	0,003	0,050	0,041	0,519
2. Promedio Factor Correlación con filtro	0,461	0,207	0,003	0,048	0,041	0,519
3. Distribución T-Student con y sin filtro	0,935	0,001	0,356	0,002	-	-
4. Proporción Área-Perímetro	0,382	0,314	0,020	0,036	0,071	0,264
5. Proporción Área-Perímetro en base a elipses	0,382	0,314	0,020	0,036	0,071	0,264
6. Variación gradientes	0,382	0,314	0,020	0,036	0,071	0,264
7. Promedio Factor Correlación con dos filtros	0,317	0,397	0,003	0,048	0,055	0,287

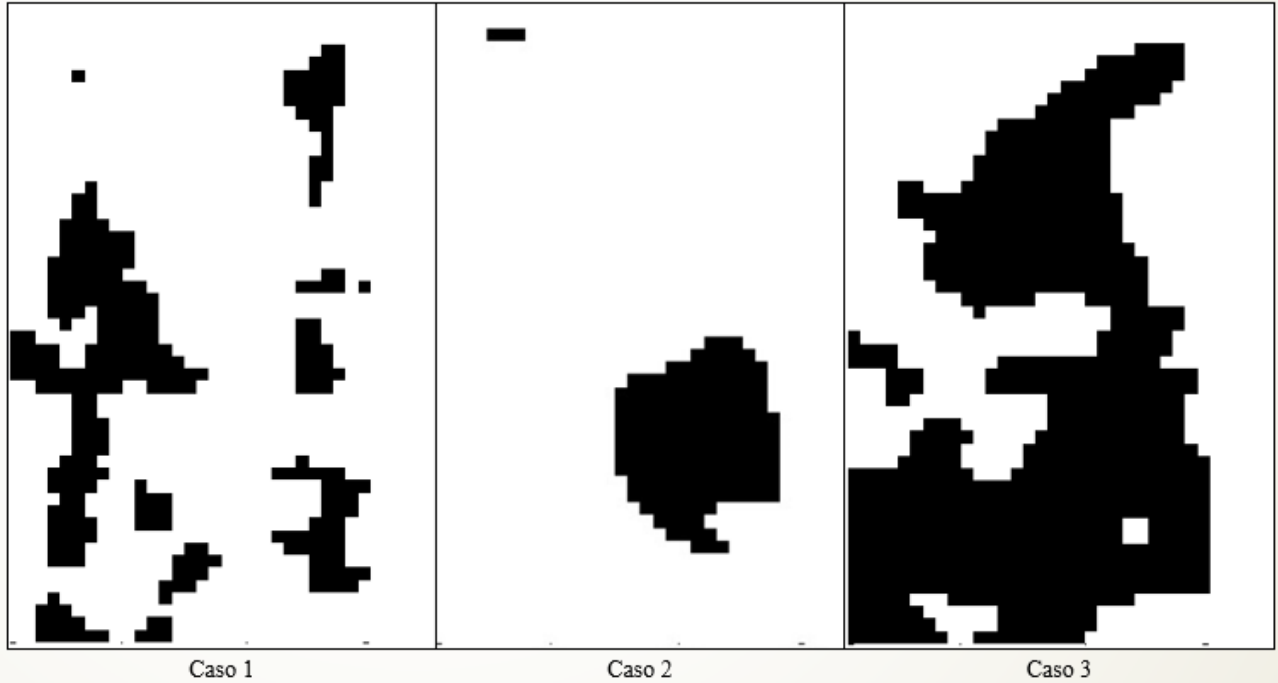


Figura 4.23: Imágenes procesadas con Ponderación de Factor Correlación para cada caso.

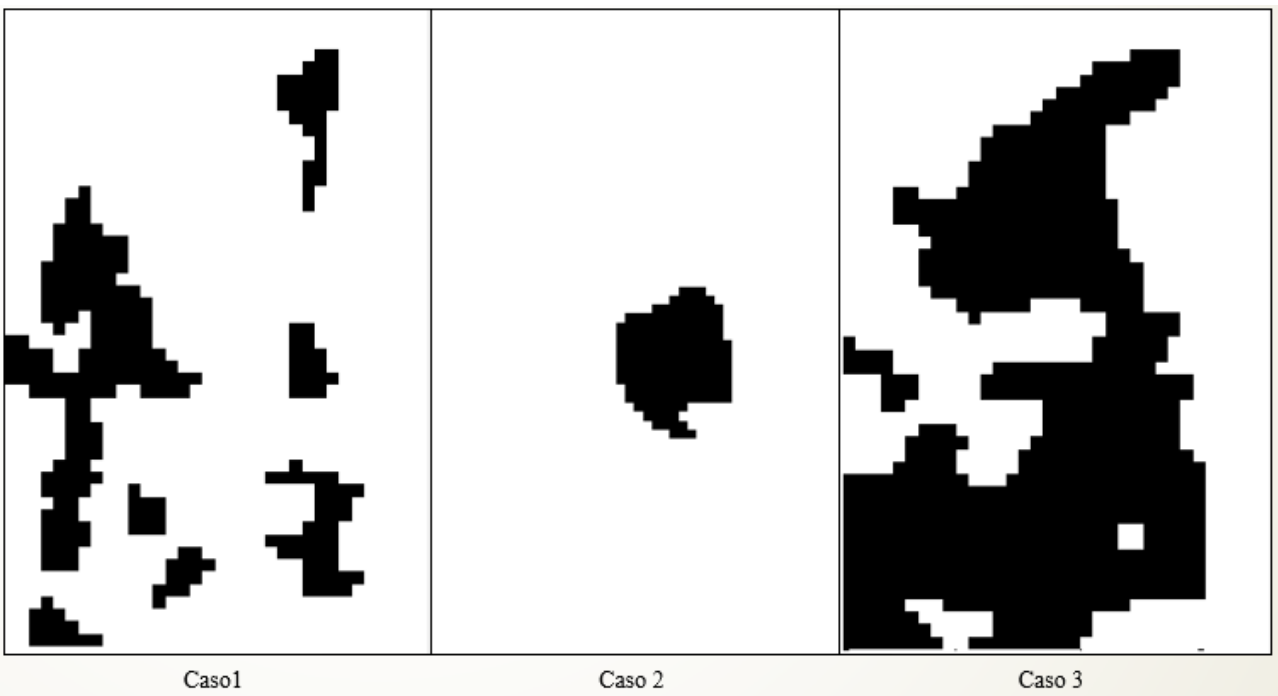


Figura 4.24: Imágenes procesadas con Ponderación de Factor Correlación con filtro para cada caso.

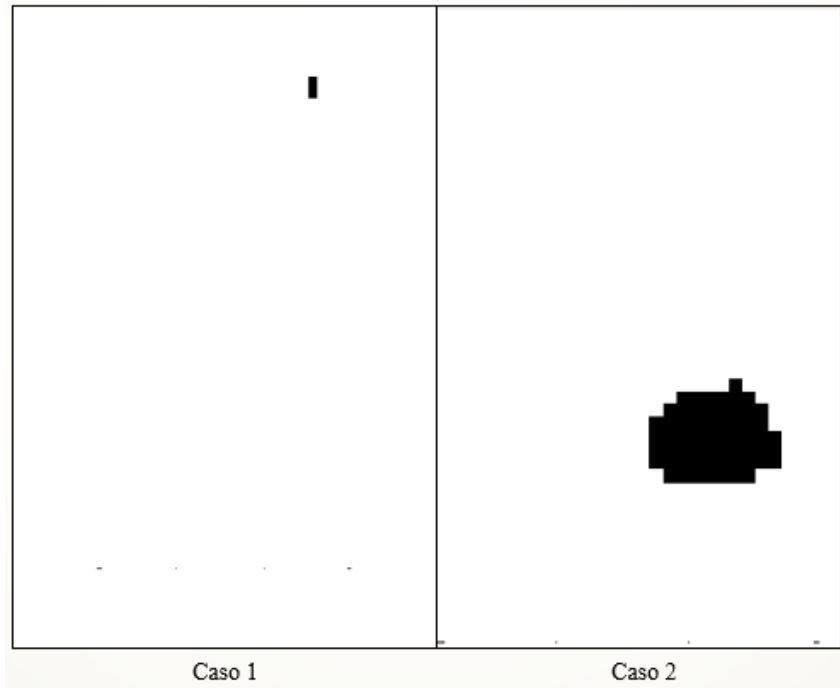


Figura 4.25: Imágenes procesadas con Método T-Student con y sin filtro para cada caso.

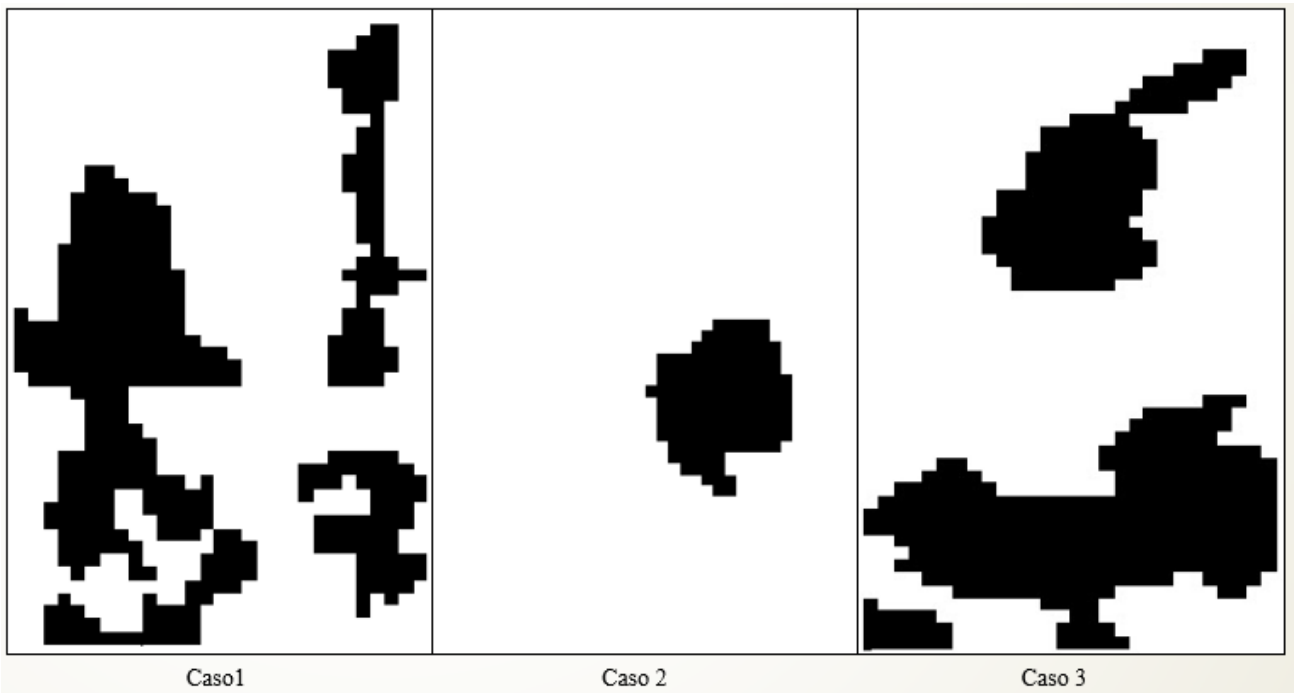


Figura 4.26: Imágenes procesadas en base a Área y Perímetro, tanto en caso normal como modelada en base a elipses para cada caso, y en base a variación de gradientes.

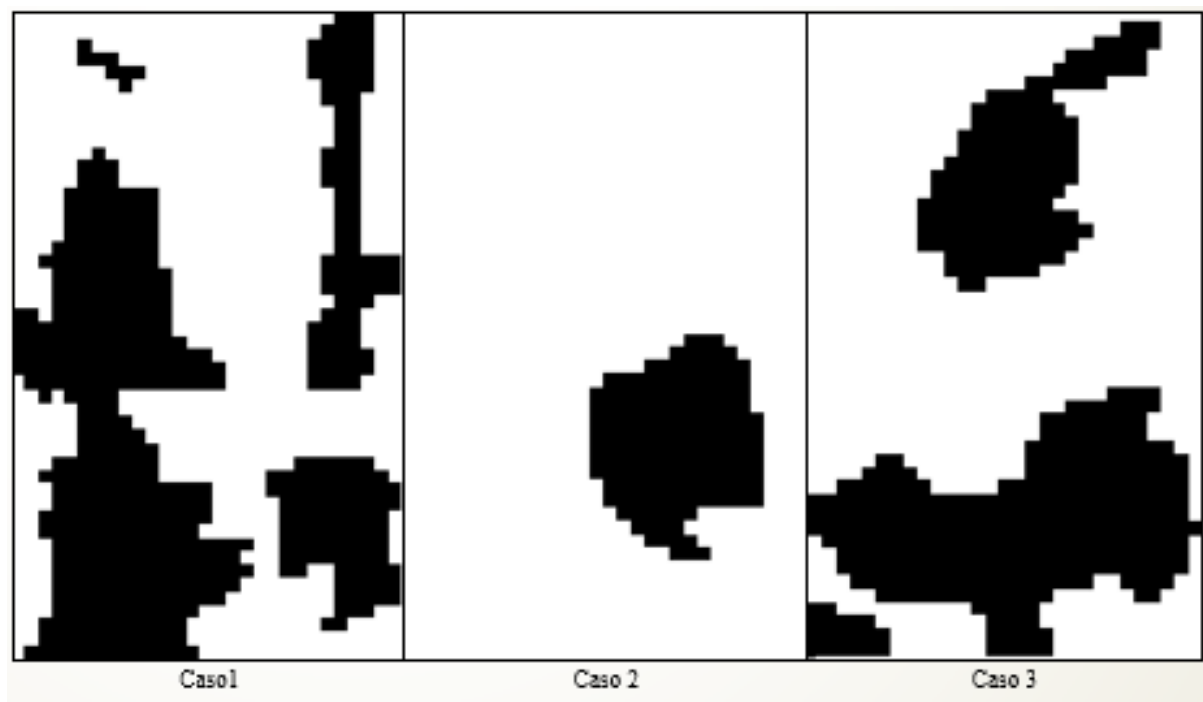


Figura 4.27: Imágenes procesadas en base a Ponderación de Factor Correlación con filtro de tamaño y en relación a área-perímetro para cada caso.

Conclusión

Se logra crear varias metodologías para el procesamiento de imágenes, de tal forma de tener una variedad de tipos de metodologías para usar. Dentro de los resultados finales se obtuvieron buenas mediciones con respecto a las imágenes simuladas numéricamente pero no con respecto a las imágenes experimentales. Esto se puede deber a que las características de las imágenes simuladas no son del todo representativas para todo tipo de imagen, o también debido a que a las imágenes experimentales no se les pueden aplicar estas metodologías debido a que sus características salen del rango estipulado o poseen características particulares.

Para el caso de Factor Correlación se encuentra un valor de 0,9437 el cual entrega un rate de FNR de 0,1 dado un rate de FPR de 0,03. Este valor es mejorado aplicando filtro de menor tamaño, con lo que se obtiene un valor de ponderación P de 0,9577 y un FNR de 0,0431 dado un FPR de 0,03. Por lo que con esta mejora se logra disminuir el FNR por lo que se logra un resultado más certero y con menos variabilidad.

En el caso del método de distribución T-Student Outliers se logra un FNR de 0,02 en el caso con filtro el cual es un poco mejor que el caso sin filtro. En este caso se obtiene un FNR menor pero un FPR mayor, por lo que de todas maneras el algoritmo es poco certero para procesar estas imágenes, además de tener una gran variabilidad.

Con respecto a los métodos de proporción entre área y perímetro modelado de forma normal y con elipses se obtienen FNR's de mayor valor, 0,053 y 0,0519 respectivamente con respecto al método de Ponderación por Factor correlación, lo cual lo hace un algoritmo menos certero pero con una variabilidad menor, lo cual lo hace un algoritmo más confiable.

El método de variación de gradiente no entrega valores muy satisfactorios, aunque los rates de FNR obtenidos fue de 0,0946 con un FPR de 0,031 no es más certero que el caso de Ponderación con Factor de Correlación. El último método entregó resultados menos satisfactorios comparado con el caso de un solo filtro de tamaño. No obstante se tiene una variación un poco menor entre los valores promedios de FNR y FPR al aplicar dos filtros en vez de uno, por lo que, el algoritmo de Factor Ponderación con filtro de tamaño y relación área y perímetro es un poco más confiable que el algoritmo base con un solo filtro.

Al momento de aplicar estas distintas metodologías en las imágenes experimentales se obtuvieron resultados no muy satisfactorios, según lo observado la gran parte de estas metodologías no lograron repetir resultados similares con respecto a las imágenes creadas numéricamente. Esto se puede deber a varias razones, primero a que al momento de entrenar o definir parámetros de los algoritmos en base a las imágenes ficticias, existen características

que las imágenes simuladas no poseen o el mismo algoritmo no toma en cuenta. Segundo se puede deber a que la forma de las imágenes simuladas y las imágenes experimentales eran diferentes, por ejemplo, las imágenes simuladas tenían ruido mucho menor y la intensidad de píxeles era notoriamente más diferente entre la zona con anomalía y sin anomalía.

Para futuros trabajos laborales se espera que se analicen las metodologías en base a las imágenes experimentales y se busque una metodología de medición de tal forma que disminuya los ruidos de estas imágenes.

Bibliografía

- [1] Joan Martí Elsa Pérez Josep Pont Erika R.E. Denton Reyer Zwiiggelaar Arnau Oliver, Jordi Freixenet. *Review of automatic mass detection and segmentation in mammographic images*. Medical Image Analysis 14. Springer London, 2010.
- [2] Xavier Lladó Meritxell Tortajada Lidia Tortajada Melcior Sentís Jordi Freixenet Reyer Zwiiggelaar Arnau Oliver, Albert Torrent. *Automatic microcalcification and cluster detection for digital and digitised mammograms*. Knowledge-Based Systems 28. 2012.
- [3] F. Grubbs. *Procedures for detecting outlying observations in samples*. Technometrics, vol. 11, no. 1. 1969.
- [4] R. Min L.M. Hu X.P. Cai H.N. Du H.D. Cheng, X.J. Shi. *Approaches for automated detection and classification of masses in mammograms*. Science Direct Pattern Recognition 39. 2006.
- [5] N Cerneaz L Tarassenko, P Hayton and M Brady. *Novelty detection for the identification of masses in mammograms*. University of Oxford.
- [6] P. Avitabile M. N. Helfrick, C. Niezrecki and T. Schmidt. *3D digital image correlation methods for full-field vibration measurement*. Mechanical Systems and Signal Processing, vol. 25, no. 3. 2011.
- [7] S.K. Gupta R.B. Dubey, M. Hanmandlu. *A comparison of two methods for the segmentation of masses in the digital mammograms*. Computerized Medical Imaging and Graphics 34. 2010.