



Universidad de Valladolid

Escuela Técnica Superior de Ingenieros de Telecomunicación

Trabajo Fin de Grado

Grado en Ingeniería de Tecnologías de Telecomunicación

MANTENIMIENTO PREDICTIVO DE UNA COSECHADORA MEDIANTE HERRAMIENTAS DE DEEP LEARNING

Presentado por:

Raúl García Hevia

Tutelado por:

Jaime Gómez Gil

Francisco Javier Gómez Gil

TITULO: MANTENIMIENTO PREDICTIVO DE UNA
COSECHADORA MEDIANTE HERRAMIENTAS DE
DEEP LEARNING

AUTOR: Raúl García Hevia

TUTOR: Jaime Gómez Gil, Departamento de Teoría de la Señal,
Comunicaciones e Ingeniería Telemática, E.T.S.I.
Telecomunicación, Universidad de Valladolid

COTUTOR: Francisco Javier Gómez Gil, Departamento de Energía
Electromecánica de la Universidad de Burgos

DEPARTAMENTO: Departamento de Teoría de la Señal, Comunicaciones e
Ingeniería Telemática

Miembros del Tribunal

PRESIDENTE: Jaime Gómez Gil

SECRETARIO: Javier Aguiar Pérez

VOCAL: Alonso Alonso Alonso

SUPLENTE: Juan Blas Prieto
Ramón de la Rosa Steinz

FECHA DE LECTURA: 8 de julio de 2024

CALIFICACIÓN:

Resumen

Monitorizar el estado de los componentes rotativos en la maquinaria agroindustrial es crucial para mantener la eficiencia operativa y prevenir fallos. Este estudio investiga el uso de Redes Neuronales Convolucionales (CNNs) para estimar el estado de los componentes rotativos basándose en señales de vibración recogidas de acelerómetros montados en el chasis de una cosechadora. Los datos se recolectaron bajo diversas condiciones operativas, reflejando diferentes estados del motor, la trilladora y el picador. El método propuesto consiste en alimentar muestras de señales de vibración en bruto en modelos CNN con diferentes configuraciones para determinar la estructura óptima de la red para la detección precisa de fallos. Se utilizaron cinco clasificadores binarios: uno para detectar si el motor está encendido o apagado, uno para el estado operativo del picador (encendido o apagado), otro para su estado de equilibrio (equilibrado o desequilibrado), y dos para la trilladora (uno para el estado operativo y otro para el estado de equilibrio). El estudio también analiza la influencia del número de capas convolucionales, el número de neuronas y el tamaño del kernel en la precisión y el tiempo de entrenamiento. Los resultados demuestran que las CNNs pueden clasificar eficazmente los estados operativos de los componentes con alta precisión, variando del 97.8% al 100% para la configuración final de la red dependiendo del clasificador, destacando su potencial para el mantenimiento predictivo en entornos agroindustriales.

PALABRAS CLAVE: Redes Neuronales Convolucionales, *deep learning*, análisis de vibraciones, monitoreo de maquinaria, componentes rotativos

Índice abreviado

Resumen	4
Índice de figuras	8
Capítulo 1: Introducción.....	10
Capítulo 2: Análisis de las tecnologías empleadas.....	15
Capítulo 3: Mantenimiento predictivo empleando señales de vibración. Descripción del problema y estado del arte.	26
Capítulo 4: Descripción del problema y toma de datos.....	31
Capítulo 5: Resultados, discusión, conclusiones y líneas futuras.....	37
Capítulo 6: Referencias	49

Índice general

Resumen	4
Índice de figuras	8
Capítulo 1: Introducción.....	10
1.1 Ámbito del proyecto	10
1.2 Objetivos.....	10
1.3 Fases y métodos.....	11
1.4 Medios disponibles	11
1.4.1 Medios <i>Software</i>	12
1.4.2 Medios <i>Hardware</i>	12
Capítulo 2: Análisis de las tecnologías empleadas.....	15
2.1 Acelerómetros piezoeléctricos.....	15
2.1.1 Tipos de acelerómetros	15
2.1.2 Efecto piezoeléctrico	18
2.2 Redes neuronales convolucionales	19
2.2.1 Aprendizaje automático y <i>deep learning</i>	19
2.2.2 Redes neuronales convolucionales	20
2.2.3 Conclusiones.....	24
Capítulo 3: Mantenimiento predictivo empleando señales de vibración. Descripción del problema y estado del arte.	26

3.1 Monitorización del funcionamiento y mantenimiento predictivo de elementos rotativos en maquinaria	27
3.1.1 Métodos basados en modelo.....	27
3.1.2 Métodos <i>data-driven</i>	28
3.2 Monitorización del funcionamiento y mantenimiento predictivo de máquinas agrícolas empleando señales de vibración.....	29
Capítulo 4: Descripción del problema y toma de datos.....	31
4.1 Descripción del problema.....	31
4.2 Toma de datos y preparación de muestras.....	34
Capítulo 5: Resultados, discusión, conclusiones y líneas futuras.....	37
5.1 Resultados.....	37
5.1.1 Tamaño del <i>batch</i> y número de <i>epochs</i>	37
5.1.2 Análisis de la precisión en función del número de capas convolucionales ...	38
5.1.3 Análisis de la precisión en función del número de neuronas por capa convolucional	40
5.1.4 Influencia del tamaño del <i>kernel</i> en la precisión	42
5.1.5 Matrices de confusión para la configuración seleccionada	43
5.2 Discusión	44
5.3 Conclusiones.....	47
5.4 Líneas futuras	47
Capítulo 6: Referencias	49

Índice de figuras

Figura 2.1: Funcionamiento de un acelerómetro capacitivo	16
Figura 2.2: Funcionamiento de un acelerómetro capacitivo	17
Figura 2.3: Estructura habitual de un acelerómetro piezoeléctrico	18
Figura 2.4: Ejemplo de cálculo de un pixel en una convolución bidimensional	22
Figura 2.5: Arquitectura típica de una CNN	23
Figura 4.1: Representación de la señal en tiempo para los estados 1 y 10	33
Figura 4.2: Representación de la señal en tiempo para los estados 4 y 7	33
Figura 4.3: Representación de la ubicación de los acelerómetros	34
Figura 5.1: Evolución de la precisión de entrenamiento según el número de epochs ...	38
Figura 5.2: Precisión máxima, mínima y media en función del número de capas y del clasificador	39
Figura 5.3: Tiempo de entrenamiento máximo, mínimo y medio en función del número de capas y del clasificador	40
Figura 5.4: Precisión media en función de la configuración de número de neuronas por capa convolucional	41
Figura 5.5: Tiempo medio en función de la configuración de número de neuronas por capa convolucional	41
Figura 5.6: Precisión media precisión en función de la configuración del tamaño de <i>kernel</i> en cada capa convolucional	42
Figura 5.7: Tiempo medio precisión en función de la configuración del tamaño de <i>kernel</i> en cada capa convolucional	42

Capítulo 1: Introducción

Capítulo 1: Introducción

En los últimos años se ha popularizado el mantenimiento predictivo para todo tipo de maquinaria gracias al desarrollo tecnológico y a la popularización del *machine learning*. Estos avances no se han reflejado en el ámbito agrícola en la misma medida que en el industrial, existiendo muy pocos estudios en esta línea. La maquinaria agrícola resulta idónea para aplicar este tipo de métodos debido al elevado coste que supone su mantenimiento y lo crítico y costoso que puede resultar para una explotación agraria el fallo de una máquina, siendo muy conveniente evitarlo antes de que suceda.

1.1 Ámbito del proyecto

En este contexto, sumado al trabajo previo realizado por el departamento en esta línea, surge la idea de realizar el presente Trabajo Fin de Grado. A lo largo de este estudio se trata de demostrar la viabilidad del uso de herramientas de *deep learning* en la detección de fallos de máquinas agrícolas a partir de señales de vibración.

No se ha de entender este trabajo como una herramienta de aplicación inmediata al ámbito agrícola, si no como un primer acercamiento al problema que demuestre la viabilidad de estos métodos de cara a futuros trabajos.

1.2 Objetivos

El principal objetivo de este Trabajo de Fin de Grado es demostrar la utilidad del *deep learning* en el mantenimiento predictivo de máquinas agrícolas. Para ello se tomarán como referencia estudios previos realizados en el ámbito industrial. Se comparará también la efectividad del método planteado con otros más tradicionales aplicados para resolver el mismo problema en estudios previos realizados por el tutor

del presente trabajo, de esta forma podremos determinar la efectividad de nuestro método.

Se plantean también los siguientes objetivos secundarios:

- Estudio de la influencia de la dirección del acelerómetro sobre el resultado de la predicción.
- Estudio de la influencia de la ubicación del acelerómetro sobre el resultado de la predicción.

1.3 Fases y métodos

De cara a la correcta realización del estudio, resulta necesario establecer previamente las fases del mismo.

1. Adquisición de los conocimientos necesarios acerca de *deep learning* de cara al desarrollo del estudio.
2. Estudio de los principales artículos acerca del uso de señales de vibración para el mantenimiento predictivo de maquinaria agrícola.
3. Estudio de los principales artículos acerca del uso de *deep learning* con señales de vibración en problemas de mantenimiento predictivo.
4. Preprocesado de los datos y creación de los sets de entrenamiento y validación de las redes neuronales.
5. Programación y prueba de distintas topologías de redes neuronales, empleando la biblioteca *Keras* de *Python*, y selección de la más efectiva.
6. Comparación con estudios anteriores y otros artículos de la bibliografía.
7. Obtención de las líneas futuras que continúen con la evolución del estudio iniciado.
8. Redacción del artículo científico que recoja el presente estudio.

1.4 Medios disponibles

En este apartado se hará una descripción de los diferentes medios, tanto *software* como *hardware*, empleados para realización del presente Trabajo de Fin de Grado.

1.4.1 Medios *Software*

1.4.1.1 Entorno de desarrollo, lenguaje de programación y bibliotecas empleadas

El entorno de desarrollo empleado es *Jupyter Notebook*® tanto para el preprocesado de la señal como para el desarrollo de las redes neuronales. Todo ello realizado en el lenguaje de programación *Python* y utilizando la biblioteca de código abierto de redes neuronales *Keras*.

1.4.1.2 Otros programas

- *Microsoft Office Word 2016* para Windows, para la escritura del Trabajo de Fin de Grado y del estudio científico.
- *Microsoft Office Excel 2016* para Windows, empleado para la representación gráfica y análisis de ciertos datos.
- *Mendeley Reference Manager*, empleado como herramienta de gestión bibliográfica.

1.4.2 Medios *Hardware*

1.4.2.1 Ordenador portátil Microsoft Surface Book 2

Para la programación y el procesamiento y análisis de los datos se ha empleado un ordenador portátil Surface Book 2 de Microsoft, con un procesador *Intel® Core™ i7-8650U CPU @ 1.90GHz 2.11 GHz*, 16 GB de memoria RAM y el sistema operativo Windows 11 Pro.

1.4.2.2 Cosechadora New Holland TC-56

Para la toma de datos se ha empleado una cosechadora New Holland TC-56 con aproximadamente 4000 horas de trabajo.

1.4.2.3 Acelerómetros

Los acelerómetros empleados para captar las señales de vibración son 4 acelerómetros uniaxiales *PCB 333B40* con un rango de aceleración de ± 10 g, una sensibilidad de 500 mV/g y un rango de frecuencias de 0,5 a 3000 Hz.

1.5 Estructura del trabajo

Comenzaremos el presente trabajo realizando una breve descripción de las tecnologías empleadas, principalmente los acelerómetros piezoeléctricos y las redes neuronales convolucionales. A continuación, en el Capítulo 3 se realizará un estudio del estado del arte, repasando los principales estudios que abordan el análisis de señales de vibración para determinar el estado de funcionamiento de diferentes tipos de máquinas, haciendo especial énfasis en aquellos que emplean herramientas de *deep learning* para procesar dichas señales. A continuación, se realizará una breve descripción del problema y de los datos empleados. Por último, en el Capítulo 5 se presentarán los resultados, que serán la traducción de un primer draft de un artículo científico que se encuentra en proceso de finalización en el momento de presentación del presente trabajo.

Capítulo 1: Análisis de las tecnologías empleadas

Capítulo 2: Análisis de las tecnologías empleadas

A lo largo del presente trabajo se han empleado dos tecnologías que tanto por su elevado protagonismo en el estudio como por su sofisticación requieren de un análisis un poco más detallado: los acelerómetros piezoeléctricos y las redes neuronales convolucionales.

2.1 Acelerómetros piezoeléctricos

Para comprender el funcionamiento y usos de los acelerómetros piezoeléctricos es importante contextualizarlos y comprender la historia y los distintos tipos de acelerómetros existentes en el mercado.

Los acelerómetros son dispositivos empleados para medir la aceleración, tanto lineal como angular. Se puede definir como un transductor que transforma energía de vibración en una señal eléctrica proporcional a la aceleración.

Las aplicaciones de los acelerómetros son muy variadas, están muy presentes tanto en nuestra vida cotidiana como en el ámbito industrial y científico. Desde el acelerómetro presente en nuestro *smartphone* hasta los presentes en los motores de los aviones para monitorizar su estado y su salud [1], existen multitud de usos diferentes con diferentes necesidades por lo que hay una amplia variedad de sensores.

2.1.1 Tipos de acelerómetros

Existe una amplia variedad de tipos de acelerómetros en función del efecto físico empleado para transformar la aceleración en energía eléctrica. A continuación, se hace un resumen de los principales tipos de acelerómetro disponibles en el mercado:

2.1.1.1 Acelerómetros piezoresistivos

Los primeros acelerómetros desarrollados en Estados Unidos en 1923 eran de este tipo. Estos sensores están equipados con material piezoresistivo, que bajo una fuerza externa se deforma y cambia su resistencia. Hoy en día principalmente se emplean aquellos fabricados con tecnología MEMS (sistemas micro electromecánicos), estos presentan ciertas características interesantes como son la respuesta DC, su funcionamiento a temperaturas comprendidas entre los $-54\text{ }^{\circ}\text{C}$ y los $121\text{ }^{\circ}\text{C}$ y los niveles de señal típicos de aproximadamente 200 mV final de escala. Se recomienda su utilización para golpes severos y se emplean en la industria automotriz (disparo de airbag, por ejemplo) o de defensa.

2.1.1.2 Acelerómetros capacitivos

En estos la aceleración se emplea para desplazar una placa móvil que se encuentra entre dos placas fijas, generando así dos condensadores que varían su capacidad en función de la posición de la placa intermedia, Figura 2.1.

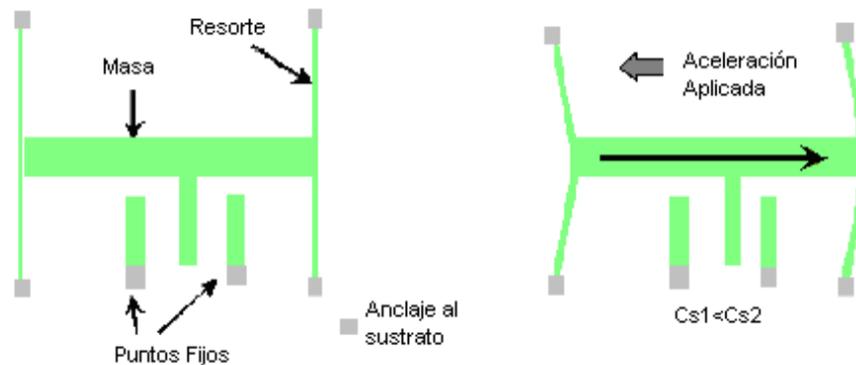


Figura 2.1: Funcionamiento de un acelerómetro capacitivo

Generan salidas de $\pm 2\text{V}$ con una gran precisión para aceleraciones de baja amplitud, funcionan en un rango térmico desde los $-54\text{ }^{\circ}\text{C}$ hasta los $121\text{ }^{\circ}\text{C}$ y tienen capacidad de respuesta en DC. Su uso está muy extendido en la industria aeroespacial y en la automovilística.

2.1.1.3 Acelerómetros de efecto Hall

En este tipo de acelerómetros una cinta conduce corriente bajo el efecto de un campo magnético no uniforme y se ve sometida a la aceleración a medir. Cuanto mayor es el desplazamiento mayor es el campo magnético al que se ve afectada la corriente

generando diferencias de potencial proporcionales a dicho movimiento debidas al efecto Hall [2]. Se utiliza en numerosos sistemas de control de tracción para determinar si las ruedas giran en sincronía, como curiosidad, es el sensor empleado para medir la velocidad de giro de las ruedas en el equipo de *Formula Student* de la Universidad de Valladolid, *Vall Racing Team*.

2.1.1.4 Acelerómetros térmicos

Este tipo de acelerómetros utiliza la transferencia de calor para medir cambios en la velocidad. Consta de un elemento calefactor rodeado de varios sensores de temperatura. En reposo el calor se difunde de forma uniforme, pero cuando se acelera cambia la distribución de dicho calor pudiendo así medirse estos cambios en la velocidad en función del calor recibido por los sensores de temperatura, Figura 2.2. Entre sus usos se encuentran la detección de la orientación y movimiento de los teléfonos móviles y la estabilización de drones.

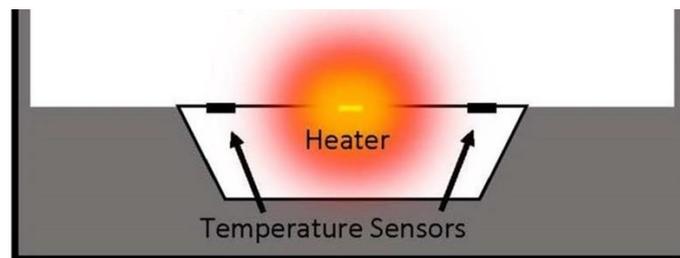


Figura 2.2: Funcionamiento de un acelerómetro capacitivo

2.1.1.5 Acelerómetros piezoeléctricos

Basan su funcionamiento en el efecto piezoeléctrico, empleando materiales piezocerámicos o cristales individuales. Destacan por el amplio rango de frecuencias en el que pueden trabajar (0,05 Hz hasta 50 KHz), su bajo peso (llegando a ser inferior a 1 gramo en algunos modelos) y su elevado rango térmico (desde temperaturas criogénicas hasta 700°C).

2.1.2 Efecto piezoeléctrico

Se denomina piezoelectricidad al fenómeno sucedido en ciertos cristales por el cual, al ser estos sometidos a tensiones mecánicas, aparece en su superficie una diferencia de potencial y cargas eléctricas. Este fenómeno se da también en el sentido opuesto, este tipo de materiales se deforman al ser sometidos a una diferencia de potencial. Se trata de un fenómeno reversible, una vez desaparece el voltaje al que está sometido el material, este recupera su forma original.

Existen dos grupos de materiales piezoeléctricos:

- Materiales de naturaleza piezoeléctrica primigenia como el cuarzo o la turmalina.
- Materiales ferroeléctricos, los cuales no presentan las características de un material piezoeléctrico hasta que no son sometidos a un proceso de polarización. Dentro de estos destaca el tantalato de litio, el nitrato de litio y la berlinita.

En concreto, para los acelerómetros basados en este efecto, se dispone un elemento piezoeléctrico de manera que, cuando vibre el sistema, una masa aplica sobre él una fuerza proporcional a dicha vibración, deformando así el material y generando, debido al efecto piezoeléctrico, la señal eléctrica proporcional a la vibración.

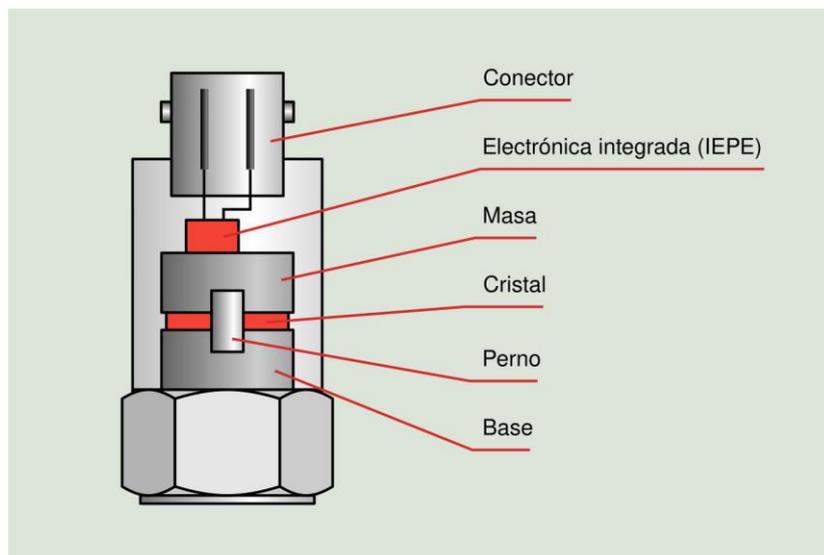


Figura 2.3: Estructura habitual de un acelerómetro piezoeléctrico

No obstante, los usos de este efecto no se limitan al campo de los acelerómetros, de hecho este tipo de materiales tienen aplicaciones muy diversas como puede ser en:

encendedores eléctricos, inyectores de combustible, altavoces, osciladores, transformadores...

2.2 Redes neuronales convolucionales

Las redes neuronales convolucionales, en adelante CNN del inglés *Convolutional Neural Network*, son un tipo de red neuronal artificial muy empleada en el campo del análisis de imagen. Para poder comprenderlas en su totalidad, es necesario antes comprender ciertos aspectos más genéricos acerca del *deep learning* y el aprendizaje automático en general.

2.2.1 Aprendizaje automático y *deep learning*

El aprendizaje automático o *machine learning* es una rama de la inteligencia artificial en la que se busca desarrollar algoritmos que aprendan a realizar ciertas funciones para las que no fueron programados específicamente de forma autónoma a partir de unos datos de entrenamiento. Existen muchas ramas del *machine learning* y cada una de ellas tiene un funcionamiento diferente, pero en general se puede concluir que lo que hace un modelo de aprendizaje automático es buscar patrones de interés en los datos para realizar la tarea para la cual está pensado.

El aprendizaje profundo o *deep learning* no es más que una categoría dentro del *machine learning*. Aunque no existe consenso en torno a su definición, la mayoría de autores concuerdan en que los algoritmos de aprendizaje profundo se caracterizan por contar con múltiples capas de procesamiento no lineal generando cada una de ellas un conjunto de características desde un nivel de abstracción más bajo a uno más alto. Se trata por tanto de redes con múltiples capas y cuyo comportamiento es complejo. Este tipo de algoritmos han ganado peso de la mano del desarrollo tecnológico debido a la gran carga computacional y al elevado volumen de datos necesarios para su entrenamiento.

2.2.1.1 Tipos de modelos de aprendizaje automático en función del grado de intervención del ser humano en su entrenamiento

En función del grado de intervención del ser humano en el entrenamiento del modelo y del conocimiento previo que se tenga sobre los datos de entrenamiento podemos diferenciar entre distintos tipos de modelos. Los principales son los siguientes:

- **Aprendizaje supervisado** – Se trata de la categoría más tradicional y en la cual el entrenamiento es más simple. El modelo aprenderá a partir de unos datos de entrada de los cuales se conoce la salida real, es decir, estos datos de entrenamiento están etiquetados. Tradicionalmente se ha dicho que se trata de un aprendizaje muy similar al humano ya que es muy similar a cuando un niño imita los comportamientos de sus padres ante situaciones similares.
- **Aprendizaje no supervisado** - En este tipo de modelos se parte de unos datos sin etiquetar y el modelo se encargará de procesarlos y buscar en ellos información o patrones que puedan ser de interés. Estos patrones, *a priori*, son desconocidos.
- **Aprendizaje semi-supervisado** – Se trata de modelos que resultan de la combinación de los dos anteriores. En estos se parte de datos etiquetados y no etiquetados y se tienen ambos tipos en cuenta de cara al entrenamiento.
- **Aprendizaje por refuerzo** – En estos modelos el sistema aprende a base de ensayo-error. El algoritmo desarrolla comportamientos que irá corrigiendo en base al *feedback* que reciba del exterior. Un ejemplo sencillo sería un algoritmo que aprende a jugar al tres en raya, en este caso el *feedback* positivo sería ganar una partida, en caso de perderla alteraría su comportamiento. En estos modelos no existe una fase puramente de entrenamiento, sino que el sistema comienza ya generando salidas y aprendiendo de su propio funcionamiento.

2.2.2 Redes neuronales convolucionales

Las redes neuronales convolucionales se basan en el Neocognitron propuesto por Kunihiko Fukushima en 1980 [3] y las mejoras implementadas en el mismo al introducir un método basado en la propagación hacia atrás o *backpropagation* por Yann LeCun en 1998 [4]. A lo largo de los últimos tiempos han experimentado una enorme popularidad en el campo del procesamiento de imagen por su enorme efectividad y su fácil implementación. Se caracterizan por estar compuestas por capas convolucionales, un tipo de capa muy concreta compuesta por neuronas convolucionales donde se realiza la operación de convolución. Aquí es donde reside la potencia de estos métodos ya que estas capas serán las encargadas de extraer características de interés de los datos de entrada que permitirán posteriormente tomar la decisión correcta a la hora de generar la salida.

2.2.2.1 Operación de convolución

La operación de convolución es ampliamente conocida y empleada en el campo del tratamiento de señal. Se trata de una operación que combina dos señales siguiendo la siguiente fórmula:

$$Y(t) = f(t) * g(t) = \int_{-\infty}^{\infty} f(u) \cdot g(t - u) du$$

En nuestro caso resulta de mayor interés la fórmula discreta ya que trabajaremos con señales digitales, esta es la siguiente:

$$Y[n] = f[n] * g[n] = \sum_{k=-\infty}^{\infty} f[k] \cdot h[n - k]$$

Por tanto, podemos decir que la convolución es una operación que genera una señal fruto de la combinación de dos señales de entrada. Tradicionalmente se ha empleado en una única dimensión, pero hoy en día en el campo del procesamiento de imagen y más en concretamente en las CNN se emplea muy habitualmente esta operación en su versión bidimensional discreta.

En este caso las entradas son dos matrices, en el campo del procesamiento de imagen una es la imagen y la otra es el filtro o *kernel*, de menor dimensión que la imagen. Tras aplicar la operación de convolución la matriz de salida se calcula barriendo la imagen con el *kernel*. El primer elemento de la matriz de salida sería fruto de situar el *kernel* de dimensión NxN sobre los NxN elementos de la esquina superior izquierda de la imagen y calculando la suma del producto de cada elemento del *kernel* por cada pixel de la imagen. Este proceso se realiza iterativamente hasta completar la imagen de entrada, se suma una influencia a la salida y se multiplica por una función de activación. En la Figura 2.4 se puede ver una iteración de este proceso.

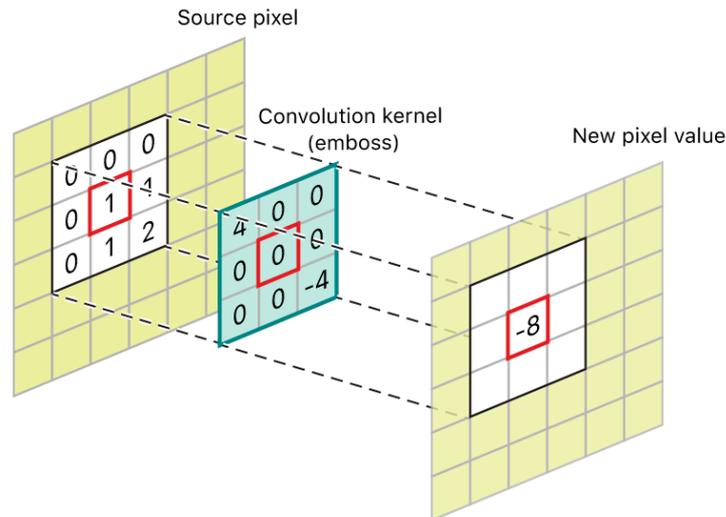


Figura 2.4: Ejemplo de cálculo de un pixel en una convolución bidimensional

2.2.2.2 Capas convolucionales

Cada capa convolucional en una CNN cuenta con numerosas neuronas convolucional. Cada una de estas neuronas aplica la convolución de la señal de entrada por un *kernel*. Este *kernel*, como se ha explicado anteriormente, sería una matriz de tamaño $N \times N$ en el caso bidimensional, un vector de tamaño N en el caso de una única dimensión. Esta matriz no es más que una matriz de pesos. Estos pesos se irán ajustando a lo largo del entrenamiento de la red neuronal para extraer características de interés. Este ajuste puede realizarse a través de diferentes mecanismos basados todos en el *backpropagation* el cual se explicará a fondo más adelante.

A la salida de cada neurona se genera una nueva matriz de tantas dimensiones como la entrada en la cual se destacan ciertas características debido al efecto del *kernel*. El objetivo del entrenamiento es que estas características destacadas sean relevantes para la función a realizar.

2.2.2.3 *Backpropagation*

El conocido como *backpropagation* es un algoritmo de *machine learning* empleado para ajustar los parámetros de una red neuronal persiguiendo el objetivo de minimizar el error cuadrático medio de sus predicciones durante su entrenamiento.

Este algoritmo se emplea en el aprendizaje supervisado. En primer lugar, se realiza una predicción utilizando datos de entrenamiento. Una vez se tiene la predicción, se

compara esta con la salida real y se calcula el error. En función de este error se ajustan los parámetros de la red (generalmente pesos y sesgos de cada nodo de la red) de tal forma que se reduzca el error. Se repite el proceso con los datos de entrenamiento hasta que la precisión en la predicción del sistema converja hasta un valor y entonces se considerará que el entrenamiento está concluido.

2.2.2.4 Arquitectura habitual de una red convolucional

Una red convolucional cuenta tradicionalmente con tres tipos de capas: convolucionales, de reducción de muestreo (también llamadas capas de *pool*) y de clasificación. A la entrada de la red se suceden concatenaciones de capas convolucionales con capas de reducción de muestreo para, a la salida, tener una neurona de clasificación (Figura 2.5).

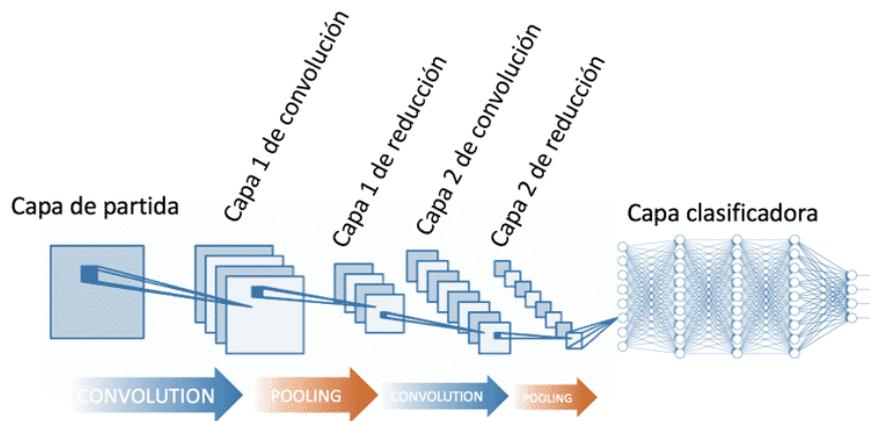


Figura 2.5: Arquitectura típica de una CNN

El funcionamiento de una capa convolucional ya ha sido explicada en el Apartado 2.2.2.2. Las capas de *pooling* son las encargadas de reducir el tamaño del vector o matriz de entrada, no se limitan a realizar la función de *subsampling*, si no que lo que hacen es realizar ciertos cálculos en conjuntos de datos de entrada para convertirlos en un único dato. Concretamente para el presente trabajo será de especial interés el conocido como *average pooling* que lo que hace es calcular la media de *n* datos para convertirlos en uno solo.

Las neuronas de clasificación se ubican al final y son las encargadas de generar la salida. Si el proceso ha finalizado, las capas convolucionales tendrán que haber sido capaces de extraer las características más esenciales de los datos de entrada y estas características se encontrarán en forma de matriz o vector. Estas neuronas de clasificación serán las encargadas clasificar estas características para generar la salida.

2.2.2.5 Aplicaciones

En los últimos años, este tipo de redes se han empleado con mucha frecuencia para la clasificación de imágenes. Uno de los campos donde han demostrado mayor potencial es en el diagnóstico de ciertas patologías a través del análisis de imágenes, durante la pandemia de COVID-19 se realizaron múltiples estudios para su diagnóstico a partir de radiografías del pulmón [5]. También han sido ampliamente empleadas este tipo de redes en reconocimiento facial y en el procesamiento de imagen para la conducción autónoma.

Aunque sus principales usos han sido en el reconocimiento de imágenes, también se han empleado redes neuronales convolucionales de 1 y 3 dimensiones. Las redes unidimensionales se emplean generalmente para trabajar con señales de audio, vibración o para procesar textos. Las de tres dimensiones suelen emplearse para señales de video al añadir a la imagen la dimensión temporal.

2.2.3 Conclusiones

Dentro del amplio campo que es el *machine learning*, para la realización del siguiente estudio se emplearán redes neuronales convolucionales. Se puede decir que las técnicas empleadas se engloban dentro del *deep learning* ya que las redes empleadas contarán con múltiples capas de procesamiento no lineal. En nuestro caso, por la naturaleza de las CNN, estas capas serán principalmente capas convolucionales y de *average pooling*, con una neurona de clasificación en al final de la red que será encargada de generar la salida.

**Capítulo 1: Mantenimiento predictivo
empleando señales de vibración. Descripción del
problema y estado del arte.**

Capítulo 3: Mantenimiento predictivo empleando señales de vibración. Descripción del problema y estado del arte.

Existen tres tipos de enfoques posibles al plantear el mantenimiento de maquinaria agrícola o industrial: el mantenimiento correctivo, el preventivo y el predictivo [6], [7].

El mantenimiento correctivo es el enfoque más convencional, se basa en la reparación de errores una vez estos ya han sucedido, generalmente habiendo limitado ya total o parcialmente estos el funcionamiento de la máquina. Es por ello que hoy en día este enfoque se tiende a aplicar sobre elementos cuyo fallo no resulta especialmente problemático, es decir, aquellos fáciles y económicos de reparar y cuyo fallo que generado sobre el funcionamiento de la máquina no resulta crítico.

El mantenimiento preventivo se realizará a lo largo de toda la vida de la máquina, comprobando periódicamente su estado y realizando los ajustes necesarios para prevenir posibles errores futuros. El objetivo de este enfoque es minimizar la necesidad del mantenimiento correctivo.

Por último, tenemos el mantenimiento predictivo, el cual se basa en la monitorización continua del funcionamiento de la máquina con el fin de detectar anomalías y establecer así un plan de mantenimiento basado en estos datos. En este contexto, el mantenimiento predictivo ha ido ganando protagonismo gracias al desarrollo tecnológico y la popularización del machine learning [8].

De entre todos los parámetros de una máquina que se pueden monitorizar, las vibraciones se han convertido en las más populares en el ámbito industrial cuando hablamos de máquinas con elementos rotatorios. Esto es debido a su elevada

efectividad, ya que se considera que generalmente este tipo de métodos son capaces de detectar alrededor del 90% de los fallos en la mayoría de los casos [9].

Para la monitorización de la máquina se pueden emplear distintas medidas y señales pero a lo largo del presente trabajo nos centraremos en aquellos métodos que utilizan señales de vibración.

3.1 Monitorización del funcionamiento y mantenimiento predictivo de elementos rotativos en maquinaria

El empleo de señales de vibración para la monitorización del estado y mantenimiento de elementos rotativos en maquinaria es algo común y ampliamente empleado a lo largo de los últimos 50 años [10].

Inicialmente únicamente existían las técnicas conocidas como monitorización de condiciones *offline*, caracterizadas por no desarrollar una medición continua de las vibraciones de la máquina, si no que se medían estas en momentos concretos para posteriormente analizar esta muestra. Este tipo de métodos resultaban caros de aplicar y poco efectivos, aunque abrieron camino en el campo del análisis de señales de vibración con estos fines.

Con el paso de los años y gracias a los avances tecnológicos, se lograron implementar las llamadas técnicas de monitorización de condiciones *online*, caracterizadas por un seguimiento continuado de las vibraciones de la máquina. Este tipo de técnicas se han popularizado enormemente al permitir realizar un seguimiento constante, de forma barata y sin necesidad de instrumentación compleja o profundos conocimientos técnicos [10].

Según la forma de realizar la predicción, podemos distinguir dos tipos de métodos: los basados en modelo y los *data-driven* [11].

3.1.1 Métodos basados en modelo

Se trata de métodos más tradicionales y requieren un profundo conocimiento previo sobre los efectos que tiene el estado de la máquina a detectar sobre las vibraciones. En estos métodos se analizan características concretas de la señal de vibración que se sabe con antelación que sirven para discriminar el estado de interés. De esta forma, si se tiene una máquina A y se quiere discriminar entre los estados 1 y 2, tendríamos que conocer, como mínimo, una característica de la señal de vibración tomada en el estado 2 que no

tenga la señal tomada en el estado 1. Si esta característica fuera, por ejemplo, un pico de la señal en una determinada frecuencia (como puede ser el caso en ciertos desequilibrios rotantes) bastaría con analizar la señal en frecuencia y buscar la presencia de ese pico para discriminar entre ambos estados. Estos métodos se basan por lo tanto en un profundo conocimiento previo de los fenómenos físicos.

3.1.2 Métodos *data-driven*

Por su parte, los métodos *data-driven*, no requieren de tanto conocimiento previo ya que hacen uso de herramientas de *machine learning*. Si bien es cierto que siempre es necesario tener cierto conocimiento de la señal de vibración y los fenómenos físicos asociados de cara a diseñar un buen sistema, estos métodos nos permiten no tener que conocer con anterioridad los efectos concretos que tiene cada estado sobre las vibraciones de la máquina. Así por ejemplo, en el mismo ejemplo anterior, nos bastaría con intuir que entre los estados 1 y 2 va a haber cierta diferencia en el dominio de la frecuencia. Sabiendo esto podríamos diseñar un modelo que, a partir de la señal en frecuencia y previo entrenamiento, sepa discriminar entre ambos estados sin llegar a saber nosotros que en una frecuencia concreta se produce el pico antes mencionado.

Dentro de estos últimos métodos cabe también distinguir dos categorías. La primera y más tradicional es aquella en la que se realizan transformaciones sobre la señal y se extraen características antes de aplicar las técnicas de *machine learning* [12]–[15]. Podríamos decir que este acercamiento se trata de un híbrido entre los métodos basados en modelo y los *data-driven* ya que, para que las características extraídas de la señal sean representativas del problema y el sistema pueda discriminar correctamente el estado de la máquina sí se requiere un cierto conocimiento previo de la señal, aunque no tanto como en el caso de los métodos basados en modelo.

La segunda categoría dentro de los modelos *data-driven* se corresponde con un enfoque más contemporáneo. En estos métodos el sistema de *machine learning* se alimenta directamente con la señal, ya sea en el dominio del tiempo o de la frecuencia. El sistema es capaz en la fase de entrenamiento de encontrar patrones en las señales que le permiten diferenciar el estado de la máquina, sin falta de previamente extraer características o realizar ningún procesamiento de la señal [8], [16]. Estos métodos son ideales por lo tanto cuando no se conoce con anterioridad las diferencias en las vibraciones existentes en los distintos estados a discriminar, resulte muy complicado o costoso conocerlas y no sea un problema conseguir un elevado volumen de datos para el

entrenamiento (ya que normalmente este tipo de redes de *machine learning* requieren de procesos de entrenamiento más largos).

3.1.2.1 Métodos *data-driven* empleando CNN

Dentro de estos últimos métodos han logrado una especial popularidad aquellos que hacen uso de CNN dada su potencia al trabajar con señales temporales, su fácil implementación y la elevada presión que han demostrado. A la hora de aplicar este tipo de redes sobre señales de vibración hay dos acercamientos posibles: convertir la señal de vibración en una matriz de 2-D y aplicar una [17]CNN 2-D [18], [19], o aplicar directamente una CNN 1-D [19], [20]. La existencia de los primeros se explica ya que, como se ha mencionado anteriormente en el Apartado 2.2.2, el principal campo de aplicación de las CNN hoy en día es en el procesamiento de imagen, por lo que se ha desarrollado un extenso conocimiento sobre este tipo de redes en su versión bidimensional.

3.2 Monitorización del funcionamiento y mantenimiento predictivo de máquinas agrícolas empleando señales de vibración

Si bien encontramos una bibliografía muy extensa acerca del mantenimiento predictivo en maquinaria industrial, los trabajos centrados en la maquinaria agrícola no son tan numerosos. No obstante, ya desde 2004 se ha probado la viabilidad del empleo de este tipo de métodos en el ámbito agrícola, concretamente en el mantenimiento predictivo de tractores empleando métodos basados en modelo [21]. En los últimos años, se han comenzado a emplear métodos *data driven* realizando transformaciones sobre la señal [22]. Dentro de este último apartado, destacan los trabajos realizados por el grupo de investigación al que pertenecen el tutor y el cotutor de este Trabajo de Fin de Grado, Jaime Gómez Gil y Francisco Javier Gómez Gil [23], [24]. Estos dos últimos trabajos nos servirán como referencia ya que se centran exactamente en la misma problemática y la misma maquinaria, uno empleando Máquinas Vector Soporte y otro empleando redes neuronales y algoritmos genéticos, consiguiendo en ambos casos precisiones en el entorno del 80%.

Capítulo 1: Descripción del problema y toma de datos

Capítulo 4: Descripción del problema y toma de datos

Teniendo ya un contexto general acerca del presente estudio, este capítulo se centrará en definir ya el problema concreto al que nos enfrentamos y la forma en la que se tomarán los datos.

4.1 Descripción del problema

El objetivo de este trabajo es el de discriminar entre 18 estados de funcionamiento, recogidos en la Tabla 4.1 de la cosechadora a partir de las señales de vibración. Para ello se emplearán 5 clasificadores binarios, cada uno de los cuales se encargará de analizar una situación de funcionamiento de la máquina: motor en altas/bajas revoluciones (en adelante clasificador 0), trilla encendida/apagada (en adelante clasificador 1), trilla equilibrada/desequilibrada (en adelante clasificador 2), picador encendido/apagado (en adelante clasificador 3) y picador equilibrado/desequilibrado (en adelante clasificador 4).

Los 5 clasificadores binarios darían lugar a 32 estados (2^5), pero hay ciertos estados que no tiene sentido analizar (si la trilla o el picador están apagados no se podrá analizar si están equilibrados/desequilibrados). De esta forma y como muestra la Tabla 4.1 se obtienen los 18 estados antes mencionados.

De estos clasificadores y en base a estudios anteriores, el que supone un menor reto en cuanto a la clasificación es el encargado de detectar el estado del motor. De hecho, las diferencias sobre la señal de vibración son apreciables a simple vista si la representamos como se aprecia en la Figura 4.1.

Estado	Estado del motor	Estado de la trilla		Estado del picador		
1	Bajas revoluciones	Off		Off		
2				On	Equilibrado	
3					Desequilibrado	
4		On	Equilibrado		Off	
5					On	Equilibrado
6						Desequilibrado
7			Desequilibrado		Off	
8					On	Equilibrado
9						Desequilibrado
10	Máximas revoluciones		Off		Off	
11					On	Equilibrado
12						Desequilibrado
13		On	Equilibrado		Off	
14					On	Equilibrado
15						Desequilibrado
16			Desequilibrado		Off	
17					On	Equilibrado
18						Desequilibrado

Tabla 4.1: Aceleración media de las señales de vibración en la banda de frecuencias 1-7 Hz para las vibraciones verticales y laterales obtenidas durante la prueba experimental ejecutada sobre un camino pavimentado, un camino sin pavimentar, una parcela cultivada y una parcela sin cultivar. Los valores numéricos son expresados en unidades de (m/s²).

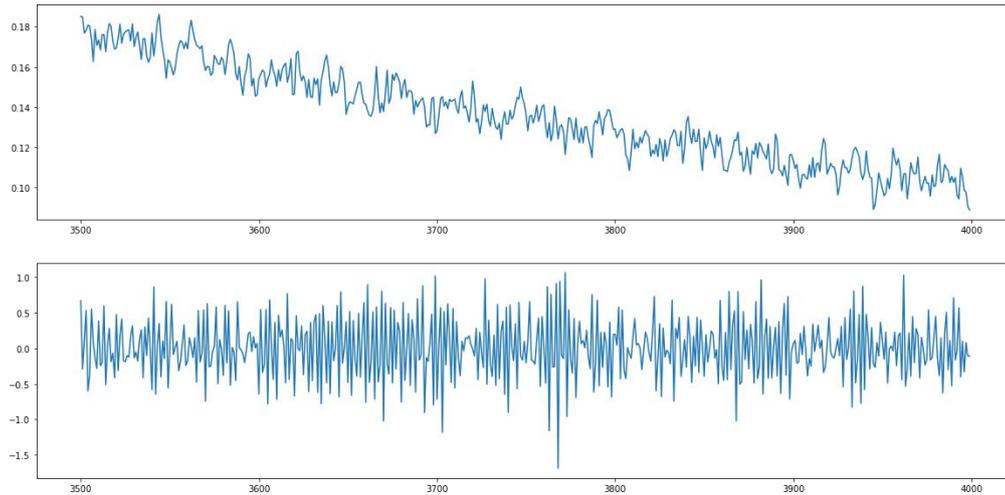


Figura 4.1: Representación de 500 puntos de la señal de vibración correspondiente a los estados 1 (motor en bajas revoluciones, imagen superior) y 10 (motor en altas revoluciones, imagen inferior)

Sin embargo, el que se espera suponga un mayor reto será el encargado de detectar los desequilibrios en el funcionamiento de la trilla.

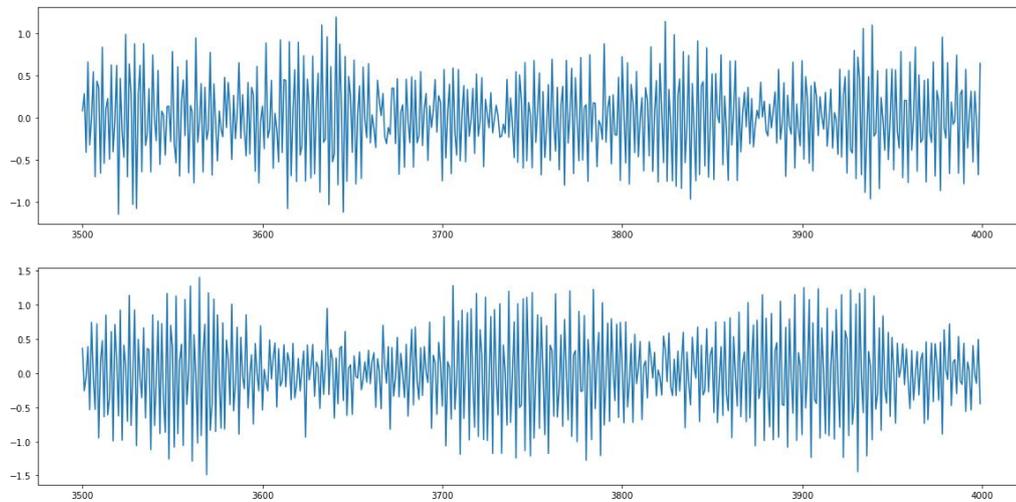


Figura 4.2: Representación de 500 puntos de la señal de vibración correspondiente a los estados 4 (trilla equilibrada, imagen superior) y 7 (trilla desequilibrada, imagen inferior)

4.2 Toma de datos y preparación de muestras

La adquisición de datos se llevó a cabo utilizando un módulo de adquisición de datos de entrada analógica de vibraciones National Instrument (NI) 9234, conectado a un portátil mediante un chasis compacto NI cDAQ (NI cDAQ-9172). El software NI Sound and Vibration Assistant gestionó la adquisición de datos.

Para la realización de este trabajo, se han tomado muestras de 102.000 puntos (unos 60 segundos de funcionamiento a una frecuencia de muestreo de 1706 Hz) para cada uno de los 18 estados recogidos en la Tabla 4.1. Las muestras se han tomado con 4 acelerómetros uniaxiales, cada uno de ellos ubicado en un punto diferente de la estructura de la cosechadora como se aprecia en la Figura 4.3. Aunque para la mayor parte de este trabajo utilizaremos únicamente un acelerómetro.

La red se alimenta con muestras de 250 puntos cada una, correspondientes a un periodo de 146,5 ms, considerando que la frecuencia de muestreo del acelerómetro es de 1706 Hz. El conjunto de entrenamiento consta de 4000 muestras para cada estado de la máquina, extraídas del 75% inicial de la señal, dejando el 25% restante para formar el conjunto de validación.

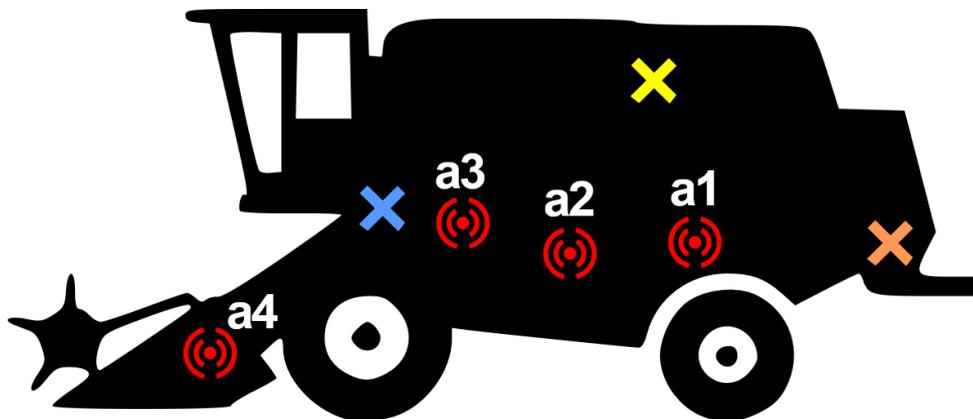


Figura 4.3: Representación de la ubicación de los acelerómetros y los puntos de interés en la máquina. Los símbolos rojos representan los acelerómetros, la cruz amarilla el motor, la azul el cilindro de la trilla y la naranja el picador.

La adquisición de señales de vibración del chasis de la cosechadora implicó la operación de los componentes rotativos mientras la máquina permanecía estacionaria y la barra de corte estaba desactivada.

Se recrearon diversas condiciones operativas en la cosechadora, dependiendo de los estados de tres componentes rotativos clave: el motor (operando a bajas RPM o RPM

máximas), el sistema de trilla (activo o inactivo) y el picador (operando o no, y en estados equilibrados o desequilibrados).

Los desequilibrios en el picador suelen surgir debido a la rotura de las cuchillas contra objetos extraños en la paja. Este fenómeno se replicó en el estudio eliminando deliberadamente una de las cuchillas del picador. Los desequilibrios en el sistema de trilla, por otro lado, comúnmente provienen del desgaste desigual de las barras debido al uso prolongado. Para simular este efecto, se fijó un peso excéntrico en el sistema de trilla.

Capítulo 1: Resultados, discusión, conclusiones y líneas futuras

Capítulo 5: Resultados, discusión, conclusiones y líneas futuras

Este capítulo se trata de una traducción al español de ciertos apartados del primer *draft* de un artículo científico, el cual se encuentra en proceso de ser terminado. Se han cambiado ciertas partes del mismo para adaptarlo a la estructura del presente trabajo. Este artículo ha sido realizado por mi, Raúl García Hevia, el profesor de la Universidad de Valladolid Jaime Gómez Gil, y los profesores de la Universidad de Burgos Francisco Javier Gómez Gil y Fernando Feijoo García.

5.1 Resultados

5.1.1 Tamaño del *batch* y número de *epochs*

Antes de iniciar las simulaciones destinadas a definir la arquitectura de nuestra red neuronal, se llevaron a cabo simulaciones preliminares con el fin de determinar el *batch size* y el número de *epochs* óptimos. La selección de estos parámetros es de suma importancia, ya que influyen significativamente en el tiempo de cómputo requerido para las simulaciones. Una determinación inadecuada del *batch size* y del número de *epochs* podría comprometer la obtención de la máxima precisión para cada tipo de red.

Con base en estudios previos, se seleccionó un *batch size* de 10 muestras por iteración para nuestra experimentación. Este *batch size* fue elegido debido a su equilibrio entre la eficiencia computacional y la estabilidad en la convergencia del modelo. A través de simulaciones iterativas, se observó que aproximadamente 40 *epochs* eran necesarias para alcanzar la precisión máxima, como se muestra en la Figura 3. Este número de *epochs* se determinó después de realizar múltiples pruebas y análisis, considerando el punto en el que la precisión del modelo se estabilizaba y dejaba de mejorar significativamente con más *epochs*.

En consecuencia, se decidió realizar 50 *epochs* en nuestras simulaciones para garantizar buenos resultados y evitar el riesgo de subentrenamiento, asegurando así que el modelo tenga suficiente exposición a los datos para aprender de manera efectiva. Este margen adicional de 10 *epochs* permite acomodar cualquier variabilidad en el proceso de entrenamiento y proporciona una mayor certeza de que el modelo alcanzará un rendimiento óptimo.

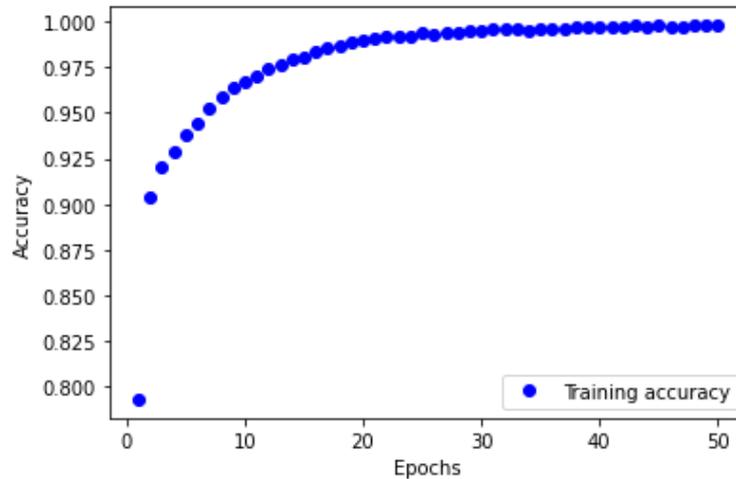


Figura 5.1: Representación de la evolución de la precisión de entrenamiento según el número de *epochs* para una de las redes simuladas.

5.1.2 Análisis de la precisión en función del número de capas convolucionales

A continuación, se presentan los resultados de precisión y tiempo de entrenamiento según el número de capas empleadas. Para lograr esto, se emplearon nueve redes diferentes para cada número de capas, variando el tamaño del *kernel* y el número de neuronas por capa convolucional.

La Tabla 5.1 representa, para cada combinación de clasificador y número de capas, la precisión mínima, máxima y promedio de 9 simulaciones realizadas variando el tamaño del *kernel* y el número de neuronas por capa convolucional, resultando en un total de 180 simulaciones. Las Figuras 5.2 y 5.3 contienen información sobre el tiempo de entrenamiento y la precisión de estas simulaciones.

	Capas	2	3	4	5
Clasificador 0	<i>Max</i>	100%	100%	100%	100%
	<i>Media</i>	100%	100%	100%	100%
	<i>Min</i>	100%	100%	100%	100%
Clasificador 1	<i>Max</i>	94,3%	97,0%	98,4%	98,4%
	<i>Media</i>	92,1%	96,4%	97,1%	97,7%
	<i>Min</i>	87,5%	95,6%	93,2%	97,1%
Clasificador 2	<i>Max</i>	93,6%	97,2%	98,2%	97,6%
	<i>Media</i>	92,3%	94,7%	97,2%	97,0%
	<i>Min</i>	90,8%	91,4%	96,5%	96,2%
Clasificador 3	<i>Max</i>	98,1%	98,6%	98,8%	98,8%
	<i>Media</i>	97,5%	98,3%	98,5%	98,5%
	<i>Min</i>	96,9%	97,9%	98,1%	98,2%
Clasificador 4	<i>Max</i>	97,7%	98,8%	98,9%	99,0%
	<i>Media</i>	96,8%	98,3%	97,9%	98,2%
	<i>Min</i>	94,7%	97,9%	96,4%	96,5%

Tabla 5.1: Tabla con los valores medio, máximo y mínimo de precisión para las primeras 180 simulaciones, en función del número de capas y del clasificador.

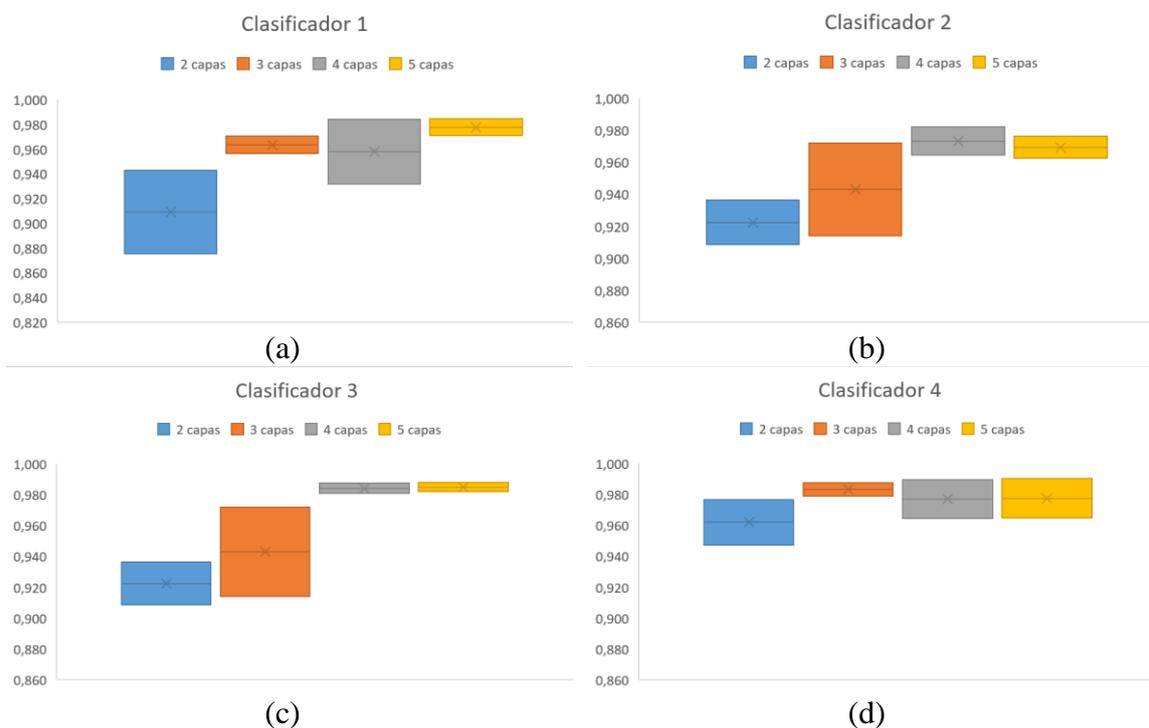


Figura 5.2: Precisión máxima, mínima y media para las primeras 180 simulaciones, en función del número de capas y del clasificador.

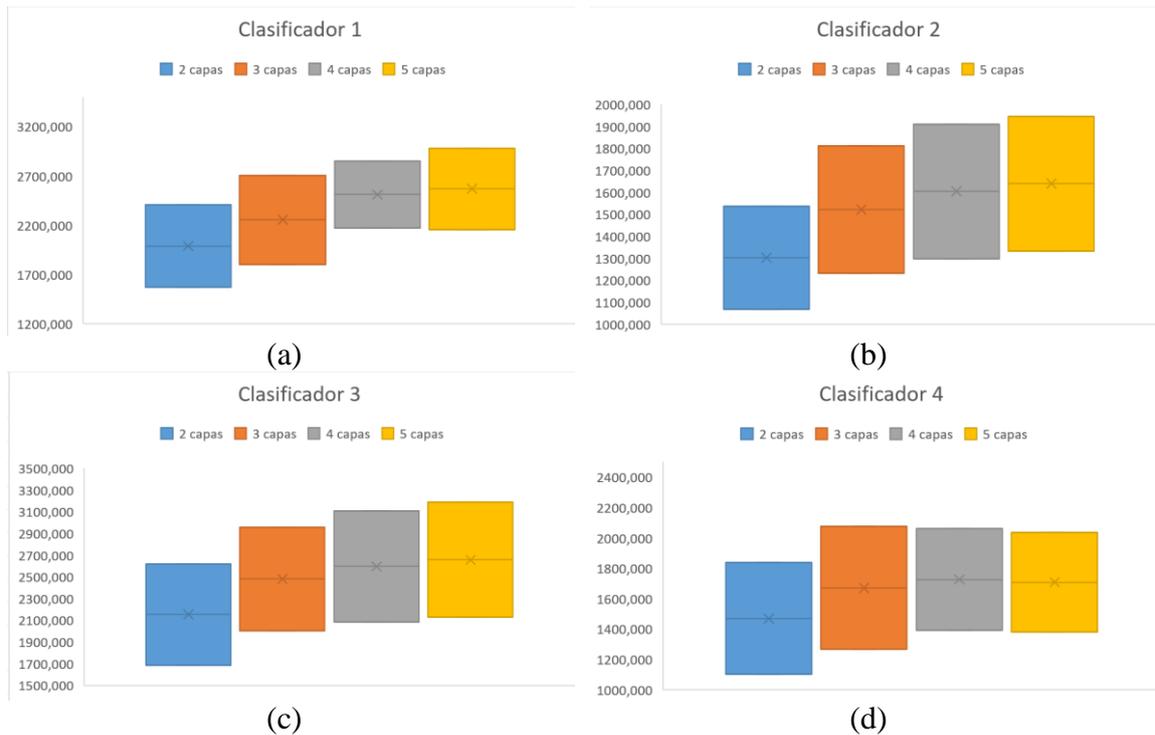


Figura 5.3: Tiempo máximo, mínimo y medio para las primeras 180 simulaciones, en función del número de capas y del clasificador.

Se observa que la precisión aumenta en general con el número de capas, pero se estabiliza entre 4 y 5 capas. Resulta complicado elegir entre modelos con 4 o 5 capas convolucionales basándose únicamente en la precisión, ya que ninguno demuestra una clara superioridad sobre el otro. Considerando los tiempos de entrenamiento, se nota que, en general, los modelos con 5 capas convolucionales requieren más tiempo de entrenamiento. Por lo tanto, los modelos con 4 capas parecen tener una ligera ventaja en este aspecto, lo que nos lleva a inclinarnos por ellos para el resto de este estudio.

5.1.3 Análisis de la precisión en función del número de neuronas por capa convolucional

La Tabla 5.2 y la Figura 5.4 y 5.5 representan los resultados promedio de precisión y tiempo de entrenamiento según el número de neuronas por capa convolucional. Para lograr esto, se realizaron simulaciones utilizando 3 combinaciones diferentes de tamaños de kernel para cada combinación del número de neuronas por capa, resultando en un total de 120 simulaciones.

Configuración	[24. 20. 16. 12.]	[32. 24. 20. 16.]	[48. 32. 24. 16.]	[48. 40. 32. 24.]	[64. 32. 24. 16.]	[64. 48. 32. 16.]	[96. 64. 32. 16.]	[128. 64. 32. 16.]
1	97,7%	97,4%	97,7%	97,5%	97,7%	95,9%	97,6%	97,6%
2	96,8%	96,2%	97,1%	95,6%	97,3%	97,2%	97,0%	97,3%
3	98,1%	98,5%	98,4%	98,3%	98,5%	98,6%	98,1%	98,5%
4	97,2%	97,6%	98,3%	98,3%	97,7%	97,6%	97,4%	97,6%

Tabla 5.2: Tabla con los valores medios de precisión en función de la configuración de número de neuronas por capa convolucional para las 120 simulaciones realizadas.

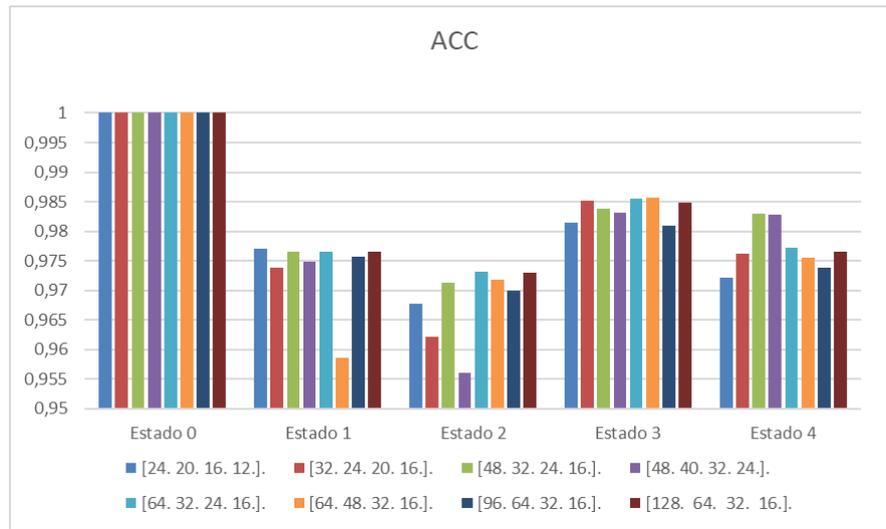


Figura 5.4: Precisión media en función de la configuración de número de neuronas por capa convolucional para las 120 simulaciones realizadas.

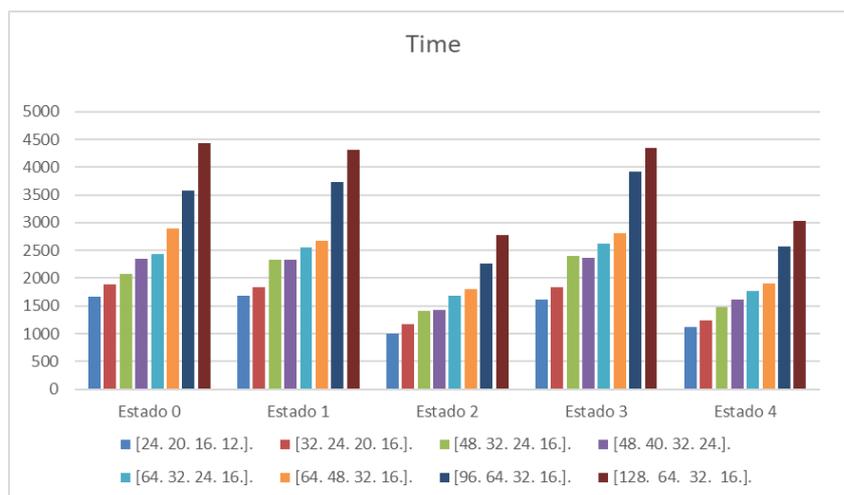


Figura 5.5: Tiempo medio en función de la configuración de número de neuronas por capa convolucional para las 120 simulaciones realizadas.

En vista de la Figura 5.5, resulta evidente que las arquitecturas con un mayor número de neuronas por capa conllevan tiempos de entrenamiento considerablemente más largos sin garantizar necesariamente una mayor precisión. Por lo tanto, seleccionar una

arquitectura de red adecuada con un número limitado de neuronas por capa convolucional será crítico para evitar tiempos de entrenamiento excesivos. En nuestro caso, optamos por la arquitectura [48, 32, 24, 16], ya que exhibe un equilibrio favorable entre precisión y tiempo de entrenamiento.

5.1.4 Influencia del tamaño del *kernel* en la precisión

La Tabla 5.3 y la Figura 5.6 representan los resultados promedio de precisión y tiempo de entrenamiento según el tamaño del *kernel* utilizado en cada una de las capas convolucionales. Para lograr esto, se realizaron simulaciones utilizando 3 combinaciones diferentes del número de neuronas por capa para cada combinación de tamaños de *kernel* empleados, resultando en un total de 120 simulaciones.

Tamaño kernel	[10. 8. 8. 7.]	[10. 9. 8. 7.]	[9. 8. 7. 6.]	[8. 8. 7. 7.]	[8. 7. 6. 5.]	[8. 6. 6. 5.]	[7. 7. 5. 5.]	[6. 5. 4. 3.]
0	100%	100%	100%	100%	100%	100%	100%	100%
1	97,7%	97,3%	97,9%	97,8%	97,5%	95,9%	97,4%	96,0%
2	97,1%	97,3%	97,8%	96,4%	96,1%	97,2%	96,5%	93,8%
3	98,4%	98,5%	98,7%	98,3%	98,2%	98,4%	98,3%	97,4%
4	98,3%	98,4%	98,5%	98,4%	98,0%	97,3%	97,7%	97,0%

Tabla 5.3: Tabla con los valores medios de precisión en función de la configuración del tamaño de *kernel* en cada capa convolucional para las 120 simulaciones realizadas.

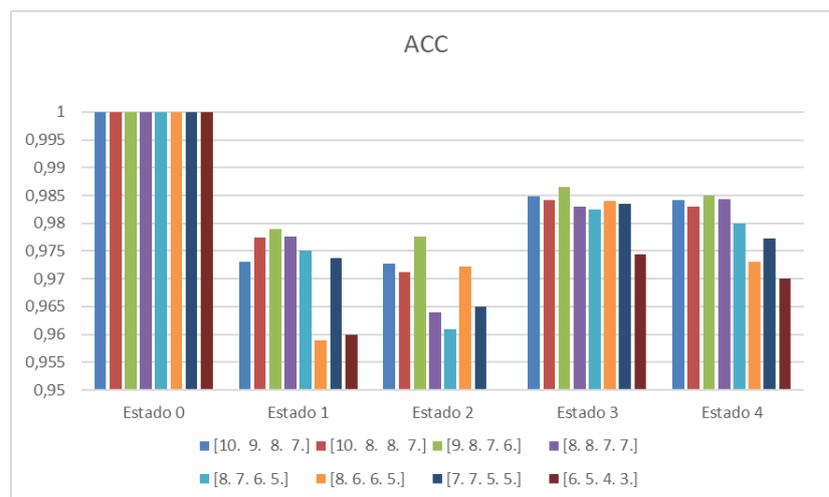


Figura 5.6: Precisión media precisión en función de la configuración del tamaño de *kernel* en cada capa convolucional para las 120 simulaciones realizadas.

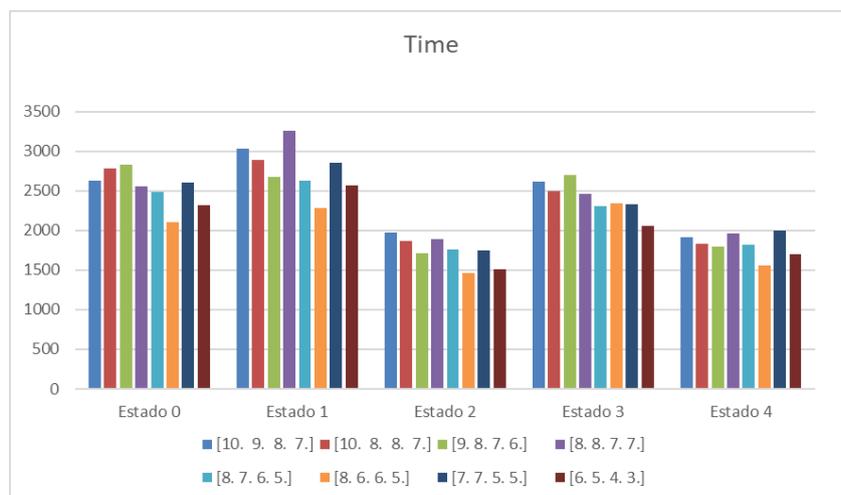


Figura 5.6: Tiempo medio precisión en función de la configuración del tamaño de *kernel* en cada capa convolucional para las 120 simulaciones realizadas.

Los *kernels* más pequeños arrojan resultados más pobres, aunque requieren menos tiempo de entrenamiento. Sin embargo, este patrón no se mantiene para los *kernels* más grandes, ya que los modelos que emplean *kernels* más grandes no necesariamente ofrecen mejores resultados. Por lo tanto, es crucial hacer elecciones informadas respecto a esta variable, buscando un tamaño de *kernel* cercano al valor óptimo. Cualquier error de cálculo, ya sea por sobredimensionar o subdimensionar el *kernel*, tendrá implicaciones para la precisión de nuestro modelo. En nuestro caso optaremos por una configuración respecto a los tamaños del *kernel* de [9,8,7,6].

5.1.5 Matrices de confusión para la configuración seleccionada

100%	0%	98,13%	4,28%	98,86%	4,67%	97,48%	0,66%	98,22%	0,94%
0%	100%	1,87%	95,72%	1,14%	95,33%	2,52%	99,34%	1,78%	99,06%

Tabla 5.4: Matrices de confusión ordenadas para los 5 clasificadores (0 a la izquierda y en orden ascendente). El valor de la esquina superior izquierda se corresponde con los verdaderos positivos, la superior derecha los falsos negativos, la inferior izquierda los falsos positivos y la inferior derecha los verdaderos negativos.

En las matrices de confusión, las asimetrías llaman especialmente la atención. Se observa que en los clasificadores asociados a la trilla, clasificadores 1 y 2, la proporción

de verdaderos positivos es significativamente mayor que la de verdaderos negativos, mientras que ocurre lo contrario en los clasificadores asociados a la picadora, clasificadores 3 y 4. Sorprendentemente, el valor de verdaderos positivos, y por tanto de falsos negativos, es muy similar en todos los clasificadores, mientras que la precisión en la detección de valores negativos varía considerablemente.

5.2 Discusión

Los resultados no deben analizarse únicamente en términos numéricos; es necesario estudiar todo el proceso para extraer conclusiones que permitan replicar el proceso al abordar problemas similares y llegar a una arquitectura de red adecuada. En este contexto, es fundamental analizar los resultados obtenidos en cada paso del proceso para detectar los elementos clave a considerar al elegir cada componente que conforma la arquitectura de la red.

El objetivo de este ejercicio es maximizar la precisión de la red mientras se minimiza el tiempo de entrenamiento. Al seleccionar el número de capas convolucionales en la red, existe una compensación entre estas dos variables. Se observa en los resultados del Apartado 5.1.2 una clara correlación positiva entre el número de capas convolucionales y el tiempo de entrenamiento de la red. Sin embargo, si se desea una precisión máxima, se debe elegir un número adecuado de capas. Los resultados indican que no hay una diferencia apreciable en la precisión al utilizar 4 o 5 capas convolucionales. No obstante, sí hay una diferencia en el tiempo de entrenamiento, por lo que se opta por redes con capas convolucionales en el resto del estudio.

En cuanto al número de neuronas por capa tratado en el Apartado 5.1.3, la situación es similar. Existe una correlación directa entre el número de neuronas y el tiempo de entrenamiento. A medida que aumenta la complejidad de la red, también crece su tiempo de entrenamiento. Sin embargo, es crucial seleccionar un número adecuado de neuronas y distribuirlas correctamente a través de las capas para maximizar la precisión mientras se mantiene el tiempo de entrenamiento en niveles manejables. Los resultados muestran una relación clara entre el número de neuronas y el tiempo de entrenamiento, pero la relación con la precisión no es tan directa. Por lo tanto, es esencial encontrar una arquitectura de red que equilibre la precisión mientras se minimiza el número total de neuronas convolucionales.

Respecto al tamaño del kernel abordado en el Apartado 5.1.4, los resultados son similares a los obtenidos para el número de neuronas, aunque parece ser una variable menos crítica, ya que muestra relaciones menos directas con el tiempo de entrenamiento y el rendimiento de la red. Es evidente que las redes con tamaños de kernel más pequeños generalmente exhiben peores resultados. Sin embargo, cuando se alcanza un nivel adecuado, los resultados se estabilizan y no se puede afirmar que tamaños de kernel más grandes conduzcan a una mayor precisión de la red. En cuanto a la relación con el tiempo de entrenamiento, parece que tamaños de kernel más grandes aumentan ligeramente el tiempo de entrenamiento, pero la relación es muy leve, mucho menor que en el caso de variar el número de neuronas. No es sorprendente que exista una relación directa, pero es menor que en el otro caso, ya que un tamaño de kernel mayor incrementa el número de operaciones a realizar para calcular cada punto de salida de cada capa convolucional, pero al mismo tiempo, provoca que la señal disminuya de tamaño más rápidamente a medida que progresa a través de las capas convolucionales de la red.

Si profundizamos en los resultados finales de precisión, podemos concluir que el método es adecuado para el problema en cuestión, ya que se han logrado niveles de precisión que van del 97,8% al 100%, variando en función del clasificador, para la configuración finalmente elegida. Se observa que los clasificadores 1 y 2, ambos asociados con la identificación del estado de la trilla, presentan los peores resultados de precisión. Esto no es sorprendente, ya que estudios previos han demostrado que las variaciones en el estado de este elemento no afectan excesivamente las vibraciones de la máquina. Sin embargo, es notable que estos dos estados muestren resultados de precisión similares cuando, en estudios anteriores, siempre fue más difícil identificar el desequilibrio que determinar si está encendido o apagado.

El presente estudio es directamente comparable con otros dentro de la misma línea de investigación llevados a cabo por nuestro equipo. Al compararlo con el estudio de 2014 empleando SVM [24], los resultados muestran una precisión del 85%, mientras que en el estudio actual se alcanza una precisión del 98.58%. De manera similar, en comparación con el estudio de 2015 que utilizó ANN [23], la precisión sigue siendo superior, con una precisión de 92.52% obtenida en ese estudio. En el último estudio, empleando un clasificador k-Nearest Neighbors (kNN) con un algoritmo selector de frecuencias Harmony Search (HS) publicado en 2024 [25], se alcanzaron precisiones del 100%, pero con muestras de 6 segundos de duración. Utilizando el método actual, se emplean muestras de 146.5 ms de duración, lo que hace impracticable el uso de métodos similares ya que tales muestras cortas pierden completamente la información sobre las frecuencias de interés primario. Una diferencia significativa en los resultados del presente estudio en comparación con todos los anteriores es que las precisiones

obtenidas para detectar el desequilibrio en la trilladora y el picador son muy similares a las obtenidas para detectar si están encendidos o apagados. A primera vista, detectar si están encendidos es mucho más simple, como corroboran estudios previos, que obtuvieron mayores precisiones para estos casos. Sin embargo, en este nuevo estudio, las precisiones para detectar desequilibrios son muy similares. Por lo tanto, podemos concluir que estos métodos pueden ser particularmente útiles para ciertas funciones donde los métodos tradicionales muestran limitaciones.

Al comparar con otros estudios que también utilizan CNN para clasificar señales de vibración basadas en el estado operativo de diferentes máquinas, se observan resultados de precisión similares. Rahimilarki logró resultados de precisión máximos alrededor del 97-98% al emplear CNN para procesar señales temporales y detectar fallas en una turbina [18]. Zhang, en 2017, procesando señales de vibración bidimensionales utilizando CNN, obtuvo niveles de precisión alrededor del 99% al detectar fallas en cojinetes [17]. Ese mismo autor, también en 2017, realizó un estudio similar añadiendo ruido, resultando en precisiones que van desde el 99.93% para una SNR de 10 dB hasta el 92.65% para una SNR de -4 dB [19]. En 2016, Ince aplicó un método muy similar para la detección de fallas en motores, obteniendo resultados de precisión del 97% [8].

Al completar nuestra investigación, surgen varias fortalezas y debilidades del presente estudio. La fortaleza de nuestra metodología radica en su utilización de datos en bruto con procesamiento de señales realizado exclusivamente en el dominio del tiempo, haciéndola altamente adaptable a problemas análogos sin necesidad de modificaciones sustanciales basadas en contextos específicos. Además, este estudio es pionero en el análisis de maquinaria agrícola, ampliando el alcance de la investigación dentro de este dominio. La consistencia de nuestra metodología se subraya por su aplicación en cinco clasificadores distintos, asegurando su generalizabilidad y robustez en contextos variados. Se lograron precisiones notablemente altas a pesar del uso de muestras con una duración de solo 146.5 milisegundos, lo que permite el análisis de problemas caracterizados por bajas frecuencias como las operaciones de la trilladora y el picador. Los métodos tradicionales habrían sido inadecuados para realizar análisis a frecuencias tan bajas con muestras de esta longitud, destacando la eficacia y superioridad de nuestro enfoque. Con la metodología propuesta sería posible responder más rápidamente a fallas repentinas, en algunos casos, reduciendo el daño causado a la máquina por dichas fallas gracias a esta rápida respuesta.

Sin embargo, las debilidades del estudio merecen también consideración. Por haber sido realizado bajo condiciones estáticas, el estudio podría estar pasando por alto el impacto del movimiento de la máquina, que puede inducir alteraciones principalmente

en frecuencias muy bajas. No obstante, se anticipa que tales cambios no afectarían significativamente los resultados. Además, otra posible debilidad es que se utilizaron datos obtenidos exclusivamente de una sola máquina, aunque con datos de 18 estados de funcionamiento distintos para cada clasificador binario, mitigando parcialmente esta limitación.

5.3 Conclusiones

Los resultados obtenidos en el estudio demuestran que (i) es posible emplear métodos de Deep Learning para detectar el estado operativo de una cosechadora a partir de señales de vibración, (ii) aplicando el método propuesto, esta tarea se puede llevar a cabo con muestras de una duración muy inferior a la necesaria para aplicar métodos tradicionales, y (iii) se establece un método replicable para definir la topología de la red para este tipo de problemas.

En cuanto a la metodología, cabe destacar que: (i) el número de capas convolucionales es crítico para la precisión de la red y el tiempo de entrenamiento; existe un punto a partir del cual aumentar este número no contribuye más a la precisión. (ii) El número de neuronas convolucionales por capa también es de vital importancia para ambas variables; en general, existe una relación directa entre este número y el tiempo de entrenamiento, aunque no con la precisión, que parece estancarse a partir de un determinado número de neuronas por capa. Por lo tanto, es importante encontrar una estructura adecuada en cuanto al número de neuronas por capa convolucional. (iii) El tamaño del *kernel* está menos relacionado con ambas variables, especialmente con la precisión, por lo que afinar correctamente este valor no es tan crítico.

5.4 Líneas futuras

Mirando hacia adelante, se presentan varias vías para futuras investigaciones. Los estudios futuros deberían explorar la aplicación de esta metodología a conjuntos de datos obtenidos de diversas máquinas, ya sean del mismo modelo o de modelos diferentes. Al analizar datos de múltiples máquinas, los investigadores pueden evaluar la generalizabilidad y escalabilidad de sus metodologías, mejorando su utilidad en un rango más amplio de contextos agrícolas. Además, aunque este estudio se centró principalmente en analizar maquinaria bajo condiciones estáticas, las investigaciones futuras deberían extender estos análisis para incluir condiciones operativas dinámicas. Finalmente, más allá del análisis de un tipo específico de maquinaria agrícola, las

investigaciones futuras deberían explorar la aplicación de la metodología propuesta a diferentes categorías de maquinaria agrícola. Las matrices de confusión, las asimetrías llaman especialmente la atención. Se observa que en los clasificadores asociados a la trilla, clasificadores 1 y 2, la proporción de verdaderos positivos es significativamente mayor que la de verdaderos negativos, mientras que ocurre lo contrario en los clasificadores asociados a la picadora, clasificadores 3 y 4. Sorprendentemente, el valor de verdaderos positivos, y por tanto de falsos negativos, es muy similar en todos los clasificadores, mientras que la precisión en la detección de valores negativos varía considerablemente.

Capítulo 6: Referencias

- [1] R. Konopinski and L. Podsedkowski, “Two planes balancing method of UAV motors using a single three-axis MEMS accelerometer,” *2022 International Conference on Unmanned Aircraft Systems, ICUAS 2022*, pp. 1474–1480, 2022, doi: 10.1109/ICUAS54217.2022.9836188.
- [2] M. Nalecz and H. Ziomecki, “A hall effect accelerometer,” *J Franklin Inst*, vol. 276, no. 1, pp. 14–25, Jul. 1963, doi: 10.1016/0016-0032(63)90306-5.
- [3] K. Fukushima, “Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position,” *Biol Cybern*, vol. 36, no. 4, pp. 193–202, Apr. 1980, doi: 10.1007/BF00344251/METRICS.
- [4] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, “Gradient-based learning applied to document recognition,” *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2323, 1998, doi: 10.1109/5.726791.
- [5] S. Kugunavar and C. J. Prabhakar, “Convolutional neural networks for the diagnosis and prognosis of the coronavirus disease pandemic,” *Vis Comput Ind Biomed Art*, vol. 4, no. 1, Dec. 2021, doi: 10.1186/S42492-021-00078-W.
- [6] R. K. Mobley, L. R. Higgins, and D. J. Wikoff, *Maintenance Engineering Handbook*, 7th ed. Mc GrawHill, 2008.
- [7] G. Sullivan, R. Pugh, A. P. Melendez, and W. D. Hunt, “Operations & Maintenance Best Practices - A Guide to Achieving Operational Efficiency (Release 3),” Aug. 2010, doi: 10.2172/1034595.
- [8] T. Ince, S. Kiranyaz, L. Eren, M. Askar, and M. Gabbouj, “Real-Time Motor Fault Detection by 1-D Convolutional Neural Networks,” *IEEE Transactions on Industrial Electronics*, vol. 63, no. 11, pp. 7067–7075, 2016, doi: 10.1109/TIE.2016.2582729.

- [9] M. Tiboni, C. Remino, R. Bussola, and C. Amici, “A Review on Vibration-Based Condition Monitoring of Rotating Machinery,” *Applied Sciences* 2022, Vol. 12, Page 972, vol. 12, no. 3, p. 972, Jan. 2022, doi: 10.3390/APP12030972.
- [10] M. Tiboni, C. Remino, R. Bussola, and C. Amici, “A Review on Vibration-Based Condition Monitoring of Rotating Machinery,” *Applied Sciences* 2022, Vol. 12, Page 972, vol. 12, no. 3, p. 972, Jan. 2022, doi: 10.3390/APP12030972.
- [11] M. H. Mohd Ghazali and W. Rahiman, “Vibration Analysis for Machine Monitoring and Diagnosis: A Systematic Review,” *Shock and Vibration*, vol. 2021, 2021, doi: 10.1155/2021/9469318.
- [12] M. Bhadane and K. I. Ramachandran, “Bearing fault identification and classification with convolutional neural network,” in *2017 International Conference on Circuit ,Power and Computing Technologies (ICCPCT)*, 2017, pp. 1–5. doi: 10.1109/ICCPCT.2017.8074401.
- [13] A. Althubaiti, F. Elasha, and J. A. Teixeira, “Fault diagnosis and health management of bearings in rotating equipment based on vibration analysis – a review,” *Journal of Vibroengineering*, vol. 24, no. 1, pp. 46–74, Nov. 2021, doi: 10.21595/JVE.2021.22100.
- [14] A. Mystkowski *et al.*, “Design and Evaluation of Low-Cost Vibration-Based Machine Monitoring System for Hay Rotary Tedder,” *Sensors* 2022, Vol. 22, Page 4072, vol. 22, no. 11, p. 4072, May 2022, doi: 10.3390/S22114072.
- [15] J. Prakash Kumar, P. S. Chauhan, and P. Prakash Pandit, “Time domain vibration analysis techniques for condition monitoring of rolling element bearing: A review,” *Mater Today Proc*, vol. 62, no. P12, pp. 6336–6340, Jan. 2022, doi: 10.1016/J.MATPR.2022.02.550.
- [16] F. Jia, Y. Lei, J. Lin, X. Zhou, and N. Lu, “Deep neural networks: A promising tool for fault characteristic mining and intelligent diagnosis of rotating machinery with massive data,” *Mech Syst Signal Process*, vol. 72–73, pp. 303–315, May 2016, doi: 10.1016/J.YMSSP.2015.10.025.
- [17] W. Zhang, G. Peng, and C. Li, “Bearings Fault Diagnosis Based on Convolutional Neural Networks with 2-D Representation of Vibration Signals as Input,” *MATEC Web of Conferences*, vol. 95, p. 13001, 2017, doi: 10.1051/mateconf/20179513001.
- [18] R. Rahimilarki, Z. Gao, N. Jin, and A. Zhang, “Convolutional neural network fault classification based on time-series analysis for benchmark wind turbine

- machine,” *Renew Energy*, vol. 185, pp. 916–931, 2022, doi: 10.1016/j.renene.2021.12.056.
- [19] W. Zhang, G. Peng, C. Li, Y. Chen, and Z. Zhang, “A New Deep Learning Model for Fault Diagnosis with Good Anti-Noise and Domain Adaptation Ability on Raw Vibration Signals,” *Sensors*, vol. 17, p. 425, 2017, doi: 10.3390/s17020425.
- [20] T. Han, C. Liu, and W. Yang, “A novel adversarial learning framework in deep convolutional neural network for intelligent diagnosis of mechanical faults,” *Knowl Based Syst*, vol. 165, pp. 474–487, 2019, doi: 10.1016/j.knosys.2018.12.019.
- [21] Roberto. Bassani, N. Pio. Belfiore, Enrico. Ciulli, and Italian association for theoretical and applied mechanics. Tribology., “Experiences of monitoring and predictive maintenance on agricultural machineries in central Italy,” *MMIV AIMETA Gentium Conventus Scientiae Tritus Roma*, vol. VOLUME UNICO, pp. 351–358, 2004, Accessed: Oct. 12, 2023. [Online]. Available: <https://iris.uniroma1.it/handle/11573/211710>
- [22] C. Mafla-Yépez, C. Castejon-Sisamon, and H. Rubio-Alonso, “Vibration Analysis in Agricultural Vehicles for Fault Detection,” *Proceedings of the XV Ibero-American Congress of Mechanical Engineering*, pp. 70–76, 2023, doi: 10.1007/978-3-031-38563-6_11.
- [23] V. Martínez-Martínez, F. J. Gomez-Gil, J. Gomez-Gil, and R. Ruiz-Gonzalez, “An Artificial Neural Network based expert system fitted with Genetic Algorithms for detecting the status of several rotary components in agro-industrial machines using a single vibration signal,” *Expert Syst Appl*, vol. 42, no. 17–18, pp. 6433–6441, Oct. 2015, doi: 10.1016/J.ESWA.2015.04.018.
- [24] R. Ruiz-Gonzalez, J. Gomez-Gil, F. J. Gomez-Gil, and V. Martínez-Martínez, “An SVM-Based Classifier for Estimating the State of Various Rotating Components in Agro-Industrial Machinery with a Vibration Signal Acquired from a Single Point on the Machine Chassis,” *Sensors 2014, Vol. 14, Pages 20713-20735*, vol. 14, no. 11, pp. 20713–20735, Nov. 2014, doi: 10.3390/S141120713.
- [25] Gomez Gil, F. J., Martinez-Martinez, V., Ruiz Gonzalez, R., & Martinez-Martinez, L. (2024). Vibration-based monitoring of agro-industrial machinery using a k-

Nearest Neighbors (kNN) classifier with a Harmony Search (HS) frequency selector algorithm. *Computers and Electronics in Agriculture*, Vol. 217.