

UNIVERSIDAD MAYOR DE SAN ANDRÉS
FACULTAD DE INGENIERÍA
CARRERA DE INGENIERÍA ELECTRÓNICA



PROYECTO DE GRADO

**“Aprendizaje fin a fin para la conducción autónoma de vehículos domésticos
usando visión artificial y redes neuronales convolucionales”**

POSTULANTE: JOSE EDUARDO LARUTA ESPEJO

TUTOR: JAVIER SANABRIA GARCIA

D.A.M.: GONZALO SAMUEL CABA MORALES

LA PAZ, AGOSTO 2018

Índice general

1. Introducción	1
1.1. Antecedentes	1
1.1.1. Sistemas de Conducción Autónoma	1
1.2. Justificación del Proyecto	4
1.2.1. Justificación académica	4
1.2.2. Justificación tecnológica	5
1.2.3. Justificación técnica	5
1.3. Análisis de la problemática y planteamiento del problema	5
1.3.1. Análisis de la problemática	5
1.3.2. Planteamiento del problema	6
1.4. Objetivos	8
1.4.1. Objetivo General	8
1.4.2. Objetivos Específicos	8
1.5. Alcance	9
1.6. Límites	9
2. Marco Teórico	11
2.1. Sistemas de Conducción Autónoma	11
2.1.1. Niveles de Autonomía	11
2.1.2. Arquitectura de un sistema de conducción autónoma	11
2.2. Visión por computador	11
2.2.1. Procesamiento de imágenes	11
2.2.2. Filtrado	11
2.3. Redes Neuronales Artificiales	11
2.3.1. Aprendizaje Automático	11
2.3.2. Aprendizaje Profundo	12
2.3.3. Redes Neuronales Convolucionales	12
2.3.4. Operación de convolución	12
2.3.5. Sistemas de Aprendizaje Fin a Fin	14
2.4. Modelo cinemático del vehículo	14
2.4.1. Ecuaciones de movimiento	15

Capítulo 1

Introducción

1.1. Antecedentes

El primer intento de desarrollo de un sistema de conducción autónomo “fin a fin” fue llevado a cabo por la Agencia de Proyectos de Investigación Avanzada en Defensa de los Estados Unidos (DARPA) con un proyecto conocido como el Vehículo Autónomo de DARPA o DAVE [1] en el cual un vehículo radio controlado a escala tenía la tarea de conducir a través de un entorno escabroso. El vehículo DAVE fue entrenado a partir de cientos de horas de conducción humana en entornos similares pero no idénticos. Los datos de entrenamiento incluyeron imágenes de dos cámaras de video y comandos de control generados por un operador humano.

Paralelamente a este esfuerzo realizado por el DARPA y debido a la limitada capacidad computacional de la época, los avances en las distintas tareas que componen la conducción autónoma se han enfocado en el tratamiento de las señales y datos provenientes de los sensores con algoritmos de procesamiento básicos llegando a crearse implementaciones efectivas basadas en un flujo de trabajo descrito a continuación.

1.1.1. Sistemas de Conducción Autónoma

Un sistema de conducción autónoma es una combinación de varios componentes o subsistemas donde las tareas de percepción, toma de decisiones y operación de un vehículo son desarrolladas por un sistema electrónico en lugar de un conductor humano.

El primer hito en el desarrollo de un sistema completamente autónomo vino con la organización del DARPA “Grand Challenge” en el cual equipos de varias universidades, institutos de investigación y empresas tuvieron que enfrentar el difícil reto de desarrollar un sistema capaz de controlar un vehículo doméstico a través de una carretera ripiada en medio del desierto de Arizona. Dentro las 2 versiones del Darpa Grand Challenge destacaron los proyectos de universidades como Stanford con el robot Stanley [2] que fue el primer vehículo en recorrer mas de 170 kilómetros en una carretera ripiada de manera completamente autónoma.

El éxito de los proyectos que participaron en el grand challenge sentó un gran precedente



Figura 1.1: Stanley, el vehículo autónomo de Stanford que ganó la competencia DARPA Grand Challenge en 2005. Fuente: stanford.edu

en el desarrollo de lo que ahora se conoce como *Self Driving Car* o vehículo autónomo. De hecho, muchos de los equipos participantes de este concurso se constituyen en la actualidad como exitosas empresas de desarrollo o coadyuvan en iniciativas privadas de gigantes de la tecnología como Google, Uber o Nissan.

Sin embargo, debido al creciente interés tanto en investigación como económico en los sistemas de conducción autónoma, la Sociedad de Ingenieros en Automoción (SAE por sus siglas en inglés) ha elaborado un estándar donde se detallan distintos aspectos concernientes. La regulación define varios niveles de autonomía en vehículos terrestres, aéreos y acuáticos yendo desde un control completamente manual, normalmente observado en vehículos completamente mecánicos, pasando por asistencias al control hasta llegar a un vehículo completamente autónomo en todas sus tareas

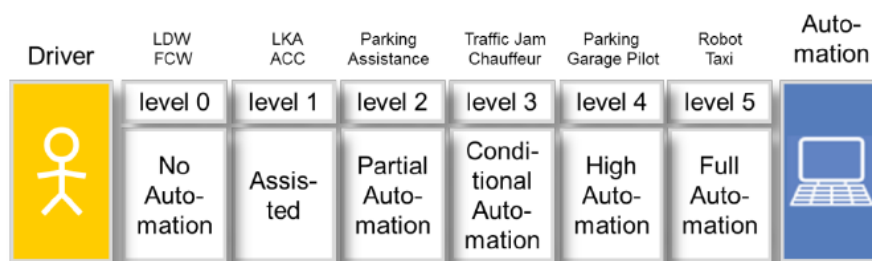


Figura 1.2: Niveles de automatización en la conducción según SAE. Fuente: researchgate

La creación de estándares y regulación ha tenido como consecuencia que, en la actualidad, existan varias iniciativas en el desarrollo de los *self Driving Cars*, siendo una de las más importantes la empresa Waymo, dependiente de Google a través de su empresa Pública Alphabet. Waymo, ha aprovechado el uso de tecnologías emergentes de sensado como el LIDAR para mejorar el mapeo y la navegación a través de algoritmos de fusión de sensores. Aparte de Alphabet, existen diversas iniciativas privadas en el desarrollo de vehículos autónomos

con fines comerciales como los Self Driving Cars de Uber, Toyota, BMW, Ford, entre otros.



Figura 1.3: El vehículo autónomo de Waymo. waymo.com

Una de las tareas más importantes dentro de un *self driving car* es la detección y mantención del carril del vehículo. Fabricantes de vehículos automotores han incluido con éxito sistemas de asistencia al conductor para la mantención del carril usando cámaras digitales y visión artificial para poder detectar la posición del automóvil con respecto al carril. Estos sistemas se consideran fundamentales en sistemas de conducción autónoma. Durante las últimas dos décadas, se han desarrollado distintos tipos de sistemas y enfoques para resolver el problema de la mantención de carril.

1.1.1.1. Arquitectura general de un sistema de conducción autónoma

En general, la arquitectura de un sistema de conducción autónoma se puede entender como la integración de varios módulos o subsistemas funcionales que operan en coordinación tal como se puede observar en la Figura(1.4).

Normalmente, este tipo de sistemas cuenta con una etapa de adquisición de datos y entrenamiento que servirá para alimentar una base de conocimiento o reglas en las que se basará el módulo de inferencia y control autónomo. Asimismo, tanto el subsistema de adquisición de datos y entrenamiento como el subsistema de inferencia y control autónomo interactúan directamente con el subsistema de control y actuación del vehículo.

Los mayores esfuerzos se han enfocado principalmente al desarrollo del subsistema de inferencia y control autónomo ya que es el que define el rendimiento de un sistema de conducción autónoma en sí. Este subsistema, a su vez, puede ser analizado como un conjunto de varios módulos que interactúan entre sí de manera secuencial, como se puede apreciar en la Figura(1.5).

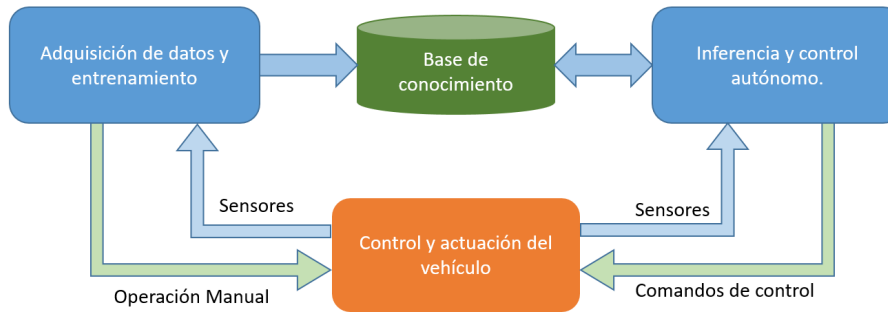


Figura 1.4: Arquitectura de un sistema de conducción autónoma. Fuente: Elaboración propia.

1.1.1.2. Sistemas de aprendizaje fin a fin

Tradicionalmente, los sistemas de aprendizaje requieren de múltiples etapas de procesamiento que interactúan entre sí, como se muestra en la Figura(1.5). Por su parte, los sistemas de aprendizaje fin a fin intentan condensar estas etapas de procesamiento y reemplazarlas usualmente con una sola red neuronal. Estos sistemas han demostrado ser altamente efectivos en contraste a los enfoques tradicionales, principalmente porque abstraen y resumen el diseño de las etapas intermedias de un sistema de aprendizaje tradicional con una sola etapa. La desventaja de los sistemas de aprendizaje fin a fin radica en la necesidad de grandes cantidades de datos de entrenamiento en comparación con los enfoques tradicionales, sin embargo, gracias a la gran disponibilidad de datos de entrenamiento y la accesibilidad de instrumentos y herramientas de adquisición de datos, esta desventaja no representa una dificultad de gran magnitud en el desarrollo de sistemas fin a fin.

Los sistemas de aprendizaje fin a fin se han explorado de manera exitosa en los últimos años, esto debido a la creciente disponibilidad de sistemas de cómputo de alta concurrencia, en especial las Unidades de Procesamiento Gráfico de propósito general o GPGPU por sus siglas en inglés. Esta disponibilidad ha logrado que se puedan entrenar redes neuronales completas en una estación de trabajo que no consume demasiada energía. Una de las empresas pioneras en GPGPU es Nvidia con su herramienta CUDA, que ha permitido el desarrollo de algoritmos de entrenamiento e inferencia para redes neuronales de manera sencilla. Es precisamente Nvidia que ha demostrado que los sistemas de aprendizaje fin a fin pueden tener éxito con el desarrollo de un prototipo y arquitectura de vehículo autónomo [3].

1.2. Justificación del Proyecto

1.2.1. Justificación académica

Desde el punto de vista académico, el proyecto se justifica en el entendido del uso de técnicas y procedimientos de ingeniería para el análisis y diseño de un sistema de aprendizaje

“fin a fin” usando redes neuronales y una plataforma para el entrenamiento y despliegue del mismo. Tales técnicas y procedimientos incluyen la definición de la arquitectura de la red neuronal, el entrenamiento y el análisis del rendimiento de la misma. Así como también el dimensionamiento de los componentes de cómputo embebido para el prototipo y la implementación de los sistemas de control electrónico de bajo nivel para el mismo. Tales técnicas y procedimientos se corresponden de manera integral con los conocimientos adquiridos a lo largo de la carrera de Ingeniería Electrónica en sus distintas asignaturas.

1.2.2. Justificación tecnológica

Dada la creciente relevancia de los sistemas autónomos en la actualidad, el proyecto se justifica desde el punto de vista tecnológico dado que se presenta la aplicación de nuevas herramientas y plataformas de software para el desarrollo de sistemas de autónomos, visión por computador y redes neuronales convolucionales, que representan áreas vigentes en la investigación tecnológica hoy en día.

El sistema desarrollado se constituirá a su vez en una plataforma de desarrollo sobre el cual se podrá extender su funcionalidad y mejorar sus resultados usando herramientas de software de fácil acceso y aprendizaje presentando la posibilidad de continuar y extender la investigación en sistemas de conducción autónoma, robótica móvil, visión por computador y aprendizaje profundo.

1.2.3. Justificación técnica

Desde el punto de vista de las técnicas aplicadas, el proyecto se justifica dado que se pretende presentar una técnica alternativa al enfoque tradicional en el desarrollo de sistemas de aprendizaje, presentando el desarrollo de un sistema de aprendizaje fin a fin, que facilitará su análisis, diseño, entrenamiento y puesta en marcha en futuros proyectos de investigación y aplicaciones en distintas áreas de la ingeniería.

La propuesta de la nueva técnica de aprendizaje fin a fin representa un avance en relación al desarrollo de sistemas tradicionales por su impacto en el requerimiento de recursos y de conocimiento específico requerido.

1.3. Análisis de la problemática y planteamiento del problema

1.3.1. Análisis de la problemática

Se han estudiado diferentes enfoques para lograr solucionar la tarea de conducción autónoma para vehículos domésticos usando sistemas de aprendizaje. Normalmente, la salida del sistema se expresa como una serie de comandos de control de aceleración y dirección del volante del vehículo. Estos comandos se pueden obtener de diversas maneras dependiendo el nivel de robustez y abstracción que el sistema requiere. Muchos sistemas se basan en la fusión

de distintos tipos de sensores y fuentes de información como ser mapas satelitales, GPS, sensores láser y cámaras. La combinación de esta información es procesada y fusionada mediante distintos algoritmos de filtrado tales como el filtro de kalman. La característica de este tipo de sistema es que se puede expresar como una serie de etapas de procesamiento mediante el cual la información fluye y se transforma, cada una de las etapas es diseñada e implementada en base a conocimiento específico y con requerimientos y limitaciones específicas de la tarea que realiza tal como se puede apreciar en la Figura(1.5).

Si bien el enfoque anteriormente mencionado ha logrado conseguir importantes avances y resultados muy prometedores, involucra un gran esfuerzo a la hora de diseñar cada una de las etapas independientemente para luego hacer que funcionen todas juntas y cumplan la tarea asignada. Este proceso usualmente requiere de un equipo de expertos que sea capaz de realizar las tareas de diseño de las etapas o módulos del sistema y el de la integración de los módulos en un solo sistema funcional. Este enfoque, pese a que ha demostrado ser una forma efectiva de trabajo para diversos problemas, tiene la desventaja de requerir muchos recursos y tiempo para poder lograr un sistema funcional.

1.3.2. Planteamiento del problema

De acuerdo a lo establecido anteriormente, se puede considerar a la etapa de inferencia y control autónomo de un sistema de conducción autónoma como un sistema de procesamiento de información que consta de varias etapas secuenciales que transforman la información de acuerdo a parámetros previamente establecidos. Se debe tener en cuenta varios aspectos concernientes tanto al diseño como a la implementación de dichos tipos de sistemas.

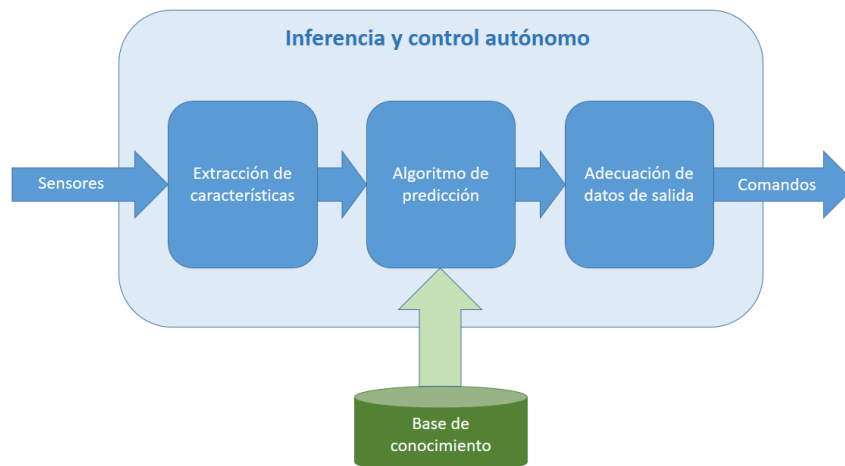


Figura 1.5: Componentes del subsistema de inferencia y control autónomo. Fuente: Elaboración propia.

En el área de visión por computadora para tareas de conducción autónoma, normalmente

se sigue el siguiente flujo en el desarrollo un sistema o prototipo:

1. **Extracción de características.** Esta etapa incluye el preprocesamiento y transformación de la imagen en un conjunto de características de distinta índole. Estas características se suelen llamar también descriptores y sirven para describir los aspectos más relevantes de la imagen para la tarea final, por ejemplo, la detección de bordes. La extracción de características también se usa para reducir la dimensionalidad inicial de la imagen a una más tratable y amigable con la capacidad de procesamiento computacional disponible. Las características o descriptores a usarse se definen manualmente por medio de conocimiento experto y se afinan de la misma manera.
2. **Algoritmo de predicción.** Esta etapa incluye típicamente un algoritmo de aprendizaje previamente entrenado con un conjunto de datos adecuado, permite realizar distintas tareas de alto nivel sobre los descriptores obtenidos de la imagen. Estas descripciones de alto nivel incluyen normalmente tareas de detección, clasificación o regresión. Los algoritmos de aprendizaje incluyen típicamente algoritmos básicos, tales como árboles de decisión, regresión lineal o máquinas de soporte vectorial ya que deben realizar la tarea de predicción en un conjunto de dimensionalidad relativamente baja (los descriptores).
3. **Adecuación de los datos de salida.** La información extraída de la anterior etapa debe procesarse para poder ser traducida a comandos de control que actúen directamente con las etapas de bajo nivel del vehículo, es decir la etapa de actuación y potencia. En esta etapa se suele incluir algún algoritmo de control realimentado para el control de motores así como también algoritmos de fusión de distintas fuentes de información para obtener finalmente una señal de comando para los actuadores.

Como se ha podido observar, el flujo de trabajo en un sistema de conducción autónomo se compone de varias etapas secuenciales que se deben realizar con conocimiento y experiencia específica en cada una de las mismas.

Por su parte, otra de las dificultades con este acercamiento, al reto de la conducción autónoma es el de la reducida flexibilidad del sistema. En otras palabras, si se quisiera modificar el sistema para agregar requerimientos o expandir la funcionalidad del mismo, se debe realizar una modificación a la etapa específica y evaluar el impacto de las modificaciones en todo el sistema en su conjunto. Esto dificulta de manera sustancial la reutilización de diversos componentes en sistemas similares.

Finalmente, la exagerada complejidad y conocimientos requeridos para poder implementar un sistema experimental de esta naturaleza hace prácticamente imposible su desarrollo por equipos de investigación pequeños o investigadores individuales. Dada la importancia y la potencialidad de los sistemas de conducción autónoma es esencial reducir esta dificultad de implementación y experimentación.

En conclusión, el desarrollo de un sistema de conducción autónoma presenta tres principales dificultades a la hora de ser abordado:

1. Conocimiento experto de cada una de las etapas involucradas en el sistema.

2. Poca flexibilidad en el diseño y la implementación del sistema una vez establecido y probado.
3. El tiempo y recursos necesarios para poder diseñar e implementar un sistema de tal naturaleza lo hace privativo para equipos de investigación pequeños o con poco presupuesto.

1.4. Objetivos

1.4.1. Objetivo General

Diseñar un sistema de aprendizaje “fin a fin” capaz de generar de comandos de control de vehículos domésticos basado en visión artificial y redes neuronales convolucionales para facilitar el diseño e implementación de sistemas de conducción autónoma.

1.4.2. Objetivos Específicos

Para alcanzar el objetivo general será necesario:

- Estudiar los aspectos concernientes al desarrollo de sistemas de conducción autónoma y sistemas de aprendizaje.
- Analizar los requerimientos de un sistema de conducción autónoma capaz identificar y mantener su carril mientras se conduce.
- Diseñar la arquitectura de un sistema de conducción autónoma en base a los requerimientos previamente establecidos.
- Diseñar el subsistema de adquisición de datos y entrenamiento para tareas de conducción autónoma.
- Diseñar el subsistema de control y actuación para la conducción autónoma de un vehículo con características similares a las de un vehículo doméstico real.
- Diseñar el subsistema de inferencia y control autónomo basado en el uso de redes neuronales convolucionales.
- Analizar los resultados del entrenamiento e implementación del subsistema de inferencia y control autónomo.
- Realizar pruebas de rendimiento y análisis comparativos en el sistema implementado.

1.5. Alcance

El presente proyecto de grado cubre los siguientes aspectos dentro de su alcance:

- El enfoque del estudio de sistemas de conducción autónomos de vehículos terrestres con el modelo de Ackermann.
- Investigación de arquitecturas y plataformas de software para el diseño y despliegue de robots móviles y tareas de conducción autónoma.
- El desarrollo del sistema se contempla en el marco de un proyecto académico y, por tanto, será implementado usando herramientas de software comúnmente utilizadas en investigación de sistemas autónomos.

El alcance detallado previamente estará acotado a su vez por una serie de supuestos.

1.6. Límites

El sistema, por su parte, contará con ciertas restricciones detalladas a continuación:

- La tarea de conducción autónoma estará enfocada exclusivamente al seguimiento y mantención del carril basado en imágenes provenientes de una cámara sin considerar el reconocimiento e interpretación de otro tipo de información como señales de tránsito, cruces e intersecciones o la presencia de peatones, ciclistas, animales y otros objetos en la ruta.
- El prototipo a escala servirá solamente para un análisis superficial de la dinámica de un vehículo automotor doméstico tomando como punto de inicio modelos matemáticos simplificados y limitaciones de rangos de trabajo dentro de dichos modelos.
- El diseño de la arquitectura de la red neuronal estará orientado a tareas de aprendizaje supervisado y aproximación de funciones y limitado por la capacidad de procesamiento disponible en el momento de la realización del presente proyecto.

Capítulo 2

Marco Teórico

2.1. Sistemas de Conducción Autónoma

-

2.1.1. Niveles de Autonomía

2.1.2. Arquitectura de un sistema de conducción autónoma

2.2. Visión por computador

-

2.2.1. Procesamiento de imágenes

2.2.2. Filtrado

2.3. Redes Neuronales Artificiales

-

2.3.1. Aprendizaje Automático

El aprendizaje automático surge de la necesidad de poder encontrar una representación útil de las ingentes cantidades de datos que se generan

2.3.1.1. Aprendizaje supervisado

Dentro del campo del aprendizaje au

2.3.1.2. Aprendizaje no supervisado**2.3.1.3. Aprendizaje por refuerzo****2.3.2. Aprendizaje Profundo****2.3.2.1. Redes neuronales profundas****2.3.2.2. Funciones de activación****2.3.2.3. Funcion de costo****2.3.2.4. Gradientes y retropropagación****2.3.2.5. Diseño de Arquitecturas****2.3.3. Redes Neuronales Convolucionales**

Las redes neuronales convolucionales son un tipo especializado de red neuronal que sirven para procesar datos de tipo "grilla" [4]. Algunos ejemplos de datos de tipo grilla que se pueden mencionar son los siguientes:

- **Series de tiempo.** Grilla de una dimensión tomados en intervalos regulares de tiempo.
- **Imágenes digitales.** Grilla de pixeles de dos o más dimensiones (Escala de grises, RGB).

Las también llamadas redes convolucionales, han demostrado un éxito impresionante en diversas aplicaciones prácticas especialmente en el campo de la visión por computador y el procesamiento de texto y lenguaje natural. El término "red neuronal convolucional" proviene del hecho de que en este tipo de redes neuronales se utiliza una operación matemática llamada **convolución**, siendo la convolución una operación lineal especializada para procesar datos de tipo grilla.

En los párrafos posteriores, se procede a describir la operación de convolución en el contexto de redes neuronales, pues, no siempre la definición de la misma corresponde con el concepto de convolución usado en distintos campos de la ciencia y la ingeniería.

2.3.4. Operación de convolución

En su forma más general, la convolución es una operación entre dos funciones reales y su definición se puede introducir usando el concepto de un promedio ponderado. Sea una función $x(t)$ dependiente del tiempo, tanto x como t son números reales; en este caso, la función x puede entenderse como una serie de medidas en un instante de tiempo t . Considérese una segunda función de ponderación $w(\tau)$ donde τ es la antigüedad de una medida. Si se aplica la función de ponderación en cada instante de tiempo, se puede obtener una nueva función definida por:

$$s(t) = \int x(\tau)w(t - \tau)d\tau \quad (2.1)$$

Esta operación es llamada la *operación de convolución* y es denotada tradicionalmente con un asterisco:

$$s(t) = (x * w)(t) \quad (2.2)$$

En el ejemplo de la ponderación, w debe ser una función de densidad de probabilidad válida, o la salida no podrá ser considerada como un promedio ponderado. Además, w también debe ser 0 para cualquier $t < 0$, esta última característica se denomina comunmente como el principio de “causalidad”. En general, la convolución está definida para cualquier función en la cual la integral anteriormente declarada esté definida y puede ser usada para otros propósitos aparte de promedios ponderados.

Hablando en términos de una red neuronal convolucional, el primer argumento (en el ejemplo, la función x) es comunmente referido como la **entrada**, y el segundo argumento (w , en el ejemplo) es referido como el **kernel**. La salida, a su vez, es normalmente referida como el **mapa de características**.

Por su parte, cuando se trata de señales digitales, como los datos en una computadora, el tiempo tiene una naturaleza discreta, es decir, que los datos estarán disponibles en intervalos regulares de tiempo. En este caso, el índice de tiempo t puede tomar solamente valores enteros y, entonces, es válido asumir que tanto x como w están definidos solamente para valores enteros de t . De este modo, se puede definir la convolución discreta:

$$s(t) = (x * w)(t) = \sum_{\tau=-\infty}^{\infty} x(\tau)w(t - \tau) \quad (2.3)$$

En el contexto de las aplicaciones de aprendizaje automático o, más específicamente, aprendizaje profundo, la entrada es usualmente un arreglo multidimensional de datos, y el kernel es usualmente un arreglo multidimensional de parámetros que se adaptan en el proceso de aprendizaje.

2.3.4.1. Procesamiento de imágenes con redes neuronales convolucionales

La operación de convolución se usa frecuentemente sobre datos con más de una dimensión. Las imágenes digitales son un perfecto ejemplo de un arreglo multidimensional de datos. Una imagen digital se representa mediante una matriz con filas y columnas, donde cada elemento se denomina pixel y contiene información acerca de la intensidad o luminancia, para una imagen en escala de grises o el nivel de color para distintos canales en una imagen a color. Si se toma el ejemplo de la imagen en escala de grises, se tiene una entrada o imagen bidimensional I con un kernel bidimensional correspondiente K :

$$S(i, j) = (I * K)(i, j) = \sum_m \sum_n I(m, n)K(i - m, j - n) \quad (2.4)$$

Dado que la convolución es conmutativa, se puede reescribir la ecuación 2.4 como:

$$S(i, j) = (K * I)(i, j) = \sum_m \sum_n I(i - m, j - n)K(m, n) \quad (2.5)$$

Frecuentemente, la última fórmula es la más utilizada en librerías de aprendizaje profundo por su sencillez en la implementación en un sistema computacional, esto, dado que existe menos variación en el rango de valores válidos de m y n .

2.3.4.2. Aprendizaje de representaciones internas

Una de las preguntas clave en la visión por computador es el cómo generar una buena y significativa representación interna de una imagen, dado que la mayor parte de la imagen corresponde con pixeles que no aportan mucha información relevante a la tarea asignada. Por ejemplo, si se quisiera detectar un rostro dentro de una imagen, normalmente se suele encontrar una representación que ayude a aislar solamente las porciones de la imagen que pueden contener el rostro, tales como la búsqueda de contornos, bordes y características típicas de un rostro. Antes de la aparición de las redes convolucionales, estas representaciones se hallaban de manera manual y gracias al conocimiento de expertos en el área del procesamiento de imágenes. La definición de características y mapas de características era comunmente conocida como la *ingeniería de características*, en la cual los expertos creaban descriptores para tareas específicas con una gran inversión de tiempo en la sintonización fina de los mismos.

En contraste con el anterior enfoque, las redes convolucionales generan sus propias representaciones internas de manera automática gracias al aprendizaje de los parámetros de cada uno de los kernels que componen las distintas capas de la red neuronal. En principio, las redes convolucionales se inspiraron en el trabajo de Hubel y Wiesel sobre la corteza visual primaria de un gato[5]. En dicho trabajo, se logró identificar células simples que respondían de manera sobresaliente a distintas orientaciones con campos receptivos locales. Éstas células receptivas simples se pueden corresponder con los kernels de convolución usados en las redes convolucionales por la sencillez y la localidad de su campo de receptividad.

Posteriormente, las redes convolucionales ganaron una gran popularidad debido a su rendimiento en tareas de clasificación de imágenes y detección y reconocimiento de objetos en imágenes. El primer hito de su capacidad para procesar imágenes de manera efectiva fue en concurso de clasificación de imágenes de ImageNet, donde el equipo de Geoffrey Hinton logró sobrepasar el mejor resultado en precisión de clasificación por un gran margen usando una arquitectura de red convolucional [6]. En este trabajo, se pudo apreciar con gran detalle las ventajas del enfoque del aprendizaje de representaciones internas en una red convolucional.

Tal como se puede apreciar en la Figura(2.1), en la primera capa convolucional, los kernels de convolución corresponden con representaciones básicas en una imagen como la búsqueda de bordes en distintas orientaciones, esto va acorde a lo establecido anteriormente en el modelo de la corteza visual de un gato. Puede decirse entonces que las redes convolucionales emulan, en cierto modo, al proceso biológico de visión en animales.

2.3.5. Sistemas de Aprendizaje Fin a Fin

2.4. Modelo cinemático del vehículo



Figura 2.1: Kernels convolucionales de tamaño $11 \times 11 \times 3$ en la primera capa convolucional.
Fuente: [6]

2.4.1. Ecuaciones de movimiento

Bibliografía

- [1] Y. LeCun, E. Cosatto, J. Ben, U. Muller, and B. Flepp, “Dave: Autonomous off-road vehicle control using end-to-end learning,” Technical Report DARPA-IPTO Final Report, Courant Institute/CBLI, <http://www.cs.nyu.edu/yann/research/dave/index.html>, Tech. Rep., 2004.
- [2] S. Thrun, M. Montemerlo, H. Dahlkamp, D. Stavens, A. Aron, J. Diebel, P. Fong, J. Gale, M. Halpenny, G. Hoffmann, K. Lau, C. Oakley, M. Palatucci, V. Pratt, P. Stang, S. Strohband, C. Dupont, L.-E. Jendrossek, C. Koelen, C. Markey, C. Rummel, J. van Niekerk, E. Jensen, P. Alessandrini, G. Bradski, B. Davies, S. Ettinger, A. Kaehler, A. Nefian, and P. Mahoney, “Stanley: The robot that won the DARPA grand challenge,” *Journal of Field Robotics*, vol. 23, no. 9, pp. 661–692, 2006.
- [3] M. Bojarski, D. Del Testa, D. Dworakowski, B. Firner, B. Flepp, P. Goyal, L. D. Jackel, M. Monfort, U. Muller, J. Zhang *et al.*, “End to end learning for self-driving cars,” *arXiv preprint arXiv:1604.07316*, 2016.
- [4] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*. MIT Press, 2016, <http://www.deeplearningbook.org>.
- [5] Y. LeCun, K. Kavukcuoglu, C. Farabet *et al.*, “Convolutional networks and applications in vision.” in *ISCAS*, vol. 2010, 2010, pp. 253–256.
- [6] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” in *Advances in neural information processing systems*, 2012, pp. 1097–1105.