



UNIVERSIDAD ESAN
FACULTAD DE INGENIERÍA
INGENIERÍA DE TECNOLOGÍAS DE INFORMACIÓN Y SISTEMAS

**Implementación de una aplicación de visión por computadora que monitorea el consumo
de carbohidratos para la prevención del síndrome metabólico en estudiantes
universitarios**

Trabajo de investigación para el curso de Trabajo de Tesis I

Nombre alumno: Marissa Fiorella López Guerrero
Asesor: Marks Calderón

Lima, 27 de mayo de 2024

Índice general

Índice de Tablas	3
1. PLANTEAMIENTO DEL PROBLEMA	4
1.1. Descripción de la Realidad Problemática	4
1.2. Formulación del Problema	6
1.2.1. Problema General	6
1.2.2. Problemas Específicos	7
1.3. Objetivos de la Investigación	7
1.3.1. Objetivo General	7
1.3.2. Objetivos Específicos	7
1.4. Hipotesis	8
1.4.1. Hipotesis General	8
1.4.2. Hipotesis Específicas	8
1.5. Justificación de la Investigación	9
1.5.1. Teórica	9
1.5.2. Práctica	9
1.5.3. Metodológica	9
1.6. Delimitación del Estudio	10
1.6.1. Espacial	10
1.6.2. Temporal	10

1.6.3. Conceptual	10
2. MARCO TEÓRICO	11
2.1. Antecedentes de la investigación	11
2.1.1. Paper 1: A review on vision-based analysis for automatic dietary assessment (2022)	11
2.1.2. Paper 2: A Comprehensive Survey of Image-Based Food Recognition and Volume Estimation Methods for Dietary Assessment (2021)	14
2.1.3. Paper 3: DeepFood: Food Image Analysis and Dietary Assessment via Deep Model (2020)	17
2.1.4. Paper 4: Image-Based Food Classification and Volume Estimation for Dietary Assessment: A Review (2020)	19
2.1.5. Paper 5: Deep Neural Networks for Image-Based Dietary Assessment (2021)	21
Anexos	25
A. Anexo I: Árbol del problema	26
B. Anexo II: Árbol del objetivos	27
C. Anexo III: Matriz de Consistencia	28
BIBLIOGRAFÍA	30

Índice de Tablas

C.1. Matriz de consistencia. Fuente: Elaboración propia	29
---	----

Capítulo 1

PLANTEAMIENTO DEL PROBLEMA

Para este trabajo de investigación se abordará como tema principal el síndrome metabólico, el cual aumenta el riesgo de enfermedades como la diabetes tipo 2. Este estudio se enfocará en estudiantes universitarios que a menudo descuidan su salud por sus actividades académicas. Para una propuesta de solución se busca utilizar las herramientas de inteligencia artificial como la visión por computadora y las redes neuronales convolucionales con el objetivo de poder clasificar los carbohidratos de los alimentos para prevenir el síndrome metabólico en la comunidad estudiantil, en especial reducir el riesgo de padecer diabetes tipo 2.

1.1. Descripción de la Realidad Problemática

El síndrome metabólico se refiere a una condición de salud seria que incrementa la probabilidad de padecer afecciones cardíacas, diabetes, episodios cerebrovasculares y enfermedades vinculadas al acumulamiento de grasa en las arterias (aterosclerosis). Para la revista cubana de Medicina Militar ([Vera-Ponce et al., 2022](#)) el síndrome metabólico es una enfermedad que se caracteriza por la ocurrencia de patologías metabólicas en un largo plazo, como la diabetes tipo 2. Los factores de riesgo conocido para el desarrollo del síndrome metabólico comprenden el sobrepeso, hipertensión, la obesidad, la resistencia a la insulina, aspectos genéticos, el avance en la edad y, sobre todo, la mala alimentación desde una edad temprana ([Lemieux & Després, 2020](#)).

El síndrome metabólico es conocido también como el Síndrome X y la propuesta para su diagnóstico fue dada por el programa Nacional de Educación sobre el Colesterol-Panel de Tratamiento de Adultos III (NCEPATP III) el cual muestran los límites que una persona debe de cumplir para ser diagnosticada por los criterios del síndrome metabólico, los cuales son: la

obesidad abdominal, niveles elevados de triglicéridos, niveles bajos de colesterol HD, hipertensión arterial y niveles elevados de glucosa en ayunas. Entre estos criterios, la diabetes tipo 2 se muestra como un componente central, debido a la resistencia a la insulina que afecta la disfunción en la regulación del azúcar en sangre, la hipertensión arterial y obesidad abdominal. La diabetes tipo 2 no solo aumenta significativamente el riesgo de desarrollar complicaciones cardiovasculares y otros problemas de salud asociados con el síndrome metabólico, sino que también refuerza la importancia de abordar la resistencia a la insulina como un factor central en la gestión y prevención.

El riesgo a desarrollar diabetes tipo 2 es de 3 a 5 veces mayor en los pacientes con síndrome metabólico a diferencia que la población general ([Dobrowolski et al., 2022](#)). Sin importar la edad en una persona, el síndrome metabólico puede afectar a un individuo en todas las etapas de la vida, incluyendo a jóvenes universitarios. Dentro del ámbito de la vida académica de un estudiante universitario, es usual que su estilo de vida tienda a subestimar la importancia de la salud metabólica. Esto suele ocurrir como consecuencia de la dedicación a los estudios y la presión académica, lo que incrementa el riesgo de desarrollar alguno de los factores de síndrome metabólico, tales como la obesidad, hipertensión y diabetes. Por lo tanto, es importante que logren comprender la importancia de incorporar una alimentación adecuada en su estilo de vida, ello con el fin de poder garantizar un futuro más saludable. Para ello el poder reconocer la importancia de identificar qué alimentos son adecuados para prevenir el síndrome metabólico mediante su clasificación es fundamental. Sin embargo, la realidad es que la mayoría de estudiantes universitarios no cuentan con el conocimiento necesario para llevar un buen seguimiento de su alimentación.

En este sentido, gracias a los avances tecnológicos, existen herramientas de inteligencia artificial como la visión por computadora y las redes neuronales convolucionales que cumplen un papel importante en ello. La visión por computadora, a través de los descriptores, ayuda a clasificar y evaluar de forma rápida las características de interés de los alimentos. También para este caso y prevenir el síndrome metabólico, nos centraremos en el contenido de carbohidratos de un alimento. En otras palabras, se calculará el contenido de glucosa a partir de los carbohidratos de la tabla nutricional y el índice glucémico.

Para el consejo Argentino sobre Seguridad de Alimentos y Nutrición ([CISAN, 2022](#)), el índice glucémico es un método que mide y clasifica los alimentos que tienen carbohidratos que elevan la glucosa en la sangre una vez consumidos.

La Organización de las Naciones Unidas para la Alimentación y la Agricultura (FAO) define lo define como la subida bajo la curva respuesta glucémica que produce al ingerir 50 gramos de hidratos de carbono o carbohidratos. Existen alimentos que tienen poca o nada de

carbohidratos como los huevos y vegetales de hojas verdes. Aunque el índice glucémico tiene algunas limitaciones como el no considerar la respuesta insulínica.

Como indica (Albornoz Lopez & Perez Rodrigo, 2012), los componentes de la dieta y el síndrome metabólico pueden relacionarse con contenidos bajos en carbohidratos y ello puede controlar la insulina, el peso y presión arterial. También menciona que, es importante el tipo de hidrato de carbono o carbohidrato que compone una dieta.

Lo que facilita la toma de decisiones sobre dietas especializadas y la prevención del síndrome metabólico. De igual manera, las redes neuronales convolucionales o CNN, la cual es adecuada para la tarea de reconocer y clasificar imágenes en función a patrones. No obstante, el uso de descriptores de características de visión por computadora y las redes neuronales convolucionales para la clasificación de componentes alimenticios en la prevención del síndrome metabólico es un área de trabajo que sigue en constante evaluación y desarrollo, es bajo este contexto que se trabajara la presente investigación, buscando clasificar de manera eficiente y certera los componentes alimenticios que puedan reducir el riesgo de que un joven universitario padezca del síndrome metabólico mediante el uso de descriptores de características de visión por computadora y CNN para mejorar su calidad de vida a largo plazo.

1.2. Formulación del Problema

Para la formulación del problema se tomará un problema general el cual plantea la posibilidad de calificar los componentes alimenticios mediante el uso de la inteligencia artificial con el fin de prevenir el síndrome metabólico de la diabetes tipo 2, obesidad, presión arterial, entre otros usando como base los carbohidratos ingeridos. Los cuales incluyen la búsqueda de una base de datos, la identificación de descriptores efectivos y una búsqueda a la mejor arquitectura de CNN para entrenar un modelo de evaluación de alimentos para una buena salud. Adicionalmente, la Matriz de consistencia (Anexo 1).

1.2.1. Problema General

¿Es posible crear una aplicación para monitorear carbohidratos aplicando visión por computadora para prevenir el síndrome metabólico en estudiantes universitarios?

1.2.2. Problemas Específicos

- ¿Existirá alguna base de datos de imágenes de alimentos etiquetados con su contenido de carbohidratos para poder entrenar un modelo que ayude a prevenir el síndrome metabólico?
- ¿Cómo medir la efectividad de la aplicación de visión por computadora que monitorea los carbohidratos para ayudar a prevenir el síndrome metabólico en estudiantes universitarios?
- ¿Cuál es la mejor arquitectura de redes neuronales convolucionales (CNN) para poder entrenar un modelo de monitoreo de carbohidratos?
- ¿Cuál es el método de preprocesamiento más adecuado para mejorar la calidad de las imágenes de alimentos para el monitoreo de carbohidratos que presentan un alto nivel de ruido?

1.3. Objetivos de la Investigación

Para la formulación del objetivo se tomará un objetivo general el cual plantea desarrollar una herramienta capaz de clasificar los componentes alimenticios y prevenga los riesgos asociados al síndrome metabólico a través de la medicina de carbohidratos. También se mostrarán tres objetivos específicos, los cuales están relacionados con los problemas específicos propuestos anteriormente.

1.3.1. Objetivo General

Desarrollar una aplicación que monitoree carbohidratos para ayudar a prevenir enfermedades asociadas con el síndrome metabólico en estudiantes universitarios aplicando visión por computadora.

1.3.2. Objetivos Específicos

- Encontrar y/o construir una base de datos de imágenes de alimentos etiquetados con su contenido de carbohidratos para el entrenamiento de un modelo que ayude a prevenir el síndrome metabólico.

- Determinar la efectividad de la aplicación que ayuda a prevenir el síndrome metabólico en estudiantes universitarios monitoreando los carbohidratos a través de visión por computadora
- Determinar la mejor arquitectura de redes neuronales convolucionales (CNN) para entrenar un modelo de monitoreo de carbohidratos.
- Determinar la mejor arquitectura de redes neuronales convolucionales (CNN) para entrenar un modelo de monitoreo de carbohidratos.

1.4. Hipotesis

Para la formulación de la hipótesis se tomará una hipótesis general la cual menciona si clasificar los componentes alimenticios ayuda a prevenir los riesgos de síndrome metabólico y tres hipótesis específicas las cuales están relacionadas con la formulación del problema y objetivos específicos.

1.4.1. Hipotesis General

Crear una aplicación que monitoree carbohidratos mediante visión por computadora ayuda a prevenir los riesgos asociados con el síndrome metabólico en los estudiantes universitarios

1.4.2. Hipotesis Específicas

- Una base de datos de imágenes de alimentos etiquetados con su contenido de carbohidratos permite entrenar un modelo que ayude a prevenir el síndrome metabólico.
- La efectividad de la aplicación de visión por computadora que monitorea los carbohidratos ayuda a prevenir el síndrome metabólico en estudiantes universitarios
- La mejor arquitectura de redes neuronales convolucionales (CNN) permite entrenar un modelo de monitoreo de carbohidratos
- Mediante la aplicación de técnicas de preprocesamiento de imágenes, se logra mejorar la precisión del modelo de monitoreo de carbohidratos al reducir la interferencia del ruido en las imágenes de alimentos.

1.5. Justificación de la Investigación

Para la justificación de la investigación se tomarán tres puntos principales los cuales mencionan sobre la importancia de medir y consumir alimenticios saludables para poder prevenir el síndrome metabólico utilizando herramientas de inteligencia artificial, mostrando resultados que beneficiarán no solo a la comunidad universitaria sino también a trabajadores y niños en desarrollo.

1.5.1. Teórica

Esta investigación se realiza con el propósito de aportar conocimiento sobre la importancia de adoptar hábitos alimenticios saludables para prevenir los riesgos asociados con el síndrome metabólico en jóvenes universitarios de ESAN. Por ello se emplearán técnicas de inteligencia artificial, como los descriptores de visión por computadora y las redes neuronales convolucionales (CNN), con el fin de clasificar si un alimento de consumo humano tiene un impacto positivo o negativo en la salud. Los resultados de este estudio no solo beneficiarán a los jóvenes universitarios, sino que también podrán ser aplicados a personas en el ámbito laboral y niños en edad escolar que presentan patrones alimenticios deficientes.

1.5.2. Práctica

Al culminar la investigación se tendrá desarrollada una herramienta que permitirá a cualquier joven universitario mejorar sus hábitos alimenticios para prevenir el síndrome metabólico mejorando así su salud a largo plazo.

1.5.3. Metodológica

La implementación de un algoritmo de clasificación de componentes alimenticios promoverá la concientización de la importancia de la alimentación, gracias a que se tendrá más conocimiento de cómo prevenir el síndrome metabólico. Asimismo, permitirá una evaluación más precisa de los alimentos consumidos por los jóvenes universitarios y su relación con los factores de riesgo del síndrome metabólico. A través de la aplicación de inteligencia artificial, se establecerá una metodología sólida para identificar patrones nutricionales y brindar información valiosa que fomente la adopción de hábitos alimenticios más saludables, contribuyendo a

la prevención de esta afección y promoviendo el bienestar en la comunidad universitaria.

1.6. Delimitación del Estudio

Esta investigación se centrará en la clasificación de los componentes alimenticios para la prevención del síndrome metabólico en los estudiantes de la universidad ESAN, haciendo uso de descriptores de características de visión por computadora y redes neuronales convolucionales. Para los resultados de la investigación, se espera que en un futuro más personas estén interesadas en cuidar su salud a largo plazo a través del consumo de alimentos para reducir alguno de los cinco criterios de riesgo asociados con el síndrome metabólico.

1.6.1. Espacial

Para la presente investigación se harán pruebas en estudiantes de la universidad ESAN

1.6.2. Temporal

Los datos que serán necesarios para probar el buen uso de la aplicación abarcará el año 2024

1.6.3. Conceptual

Esta investigación se centrará exclusivamente en la clasificación de componentes alimenticios para la prevención del síndrome metabólico midiéndolo a través de carbohidratos, utilizando descriptores de características de visión por computadora y redes neuronales convolucionales.

Capítulo 2

MARCO TEÓRICO

En esta sección se abordará los antecedentes de la investigación relacionadas con herramientas de inteligencia artificial como la visión por computadora, también se incluirán las metodologías y resultados por cada artículo. Adicional, las bases teóricas y el marco conceptual.

2.1. Antecedentes de la investigación

Los siguientes antecedentes incluyen el resumen, la metodología y los resultados de papers de revistas y páginas importantes como National Institutes of Health (NIH) y IEEE Xplore Digital Library

2.1.1. Paper 1: A review on vision-based analysis for automatic dietary assessment (2022)

Los autores son Wang, W., Min, W., Li, T., Dong, X. & Li, H. de las universidades de Beijing en diferentes carreras, ([Wang et al., 2022](#))

Resumen

Tener una dieta saludable evita problemas de salud a largo plazo, pero algunos métodos para su evaluación suelen ser poco confiables, esto ha ido cambiando con la ayuda de la inteligencia artificial, este artículo presenta una arquitectura de Evaluación Dietética Basada en Visión (VBDA) que toma imágenes de una comida como entrada para luego usar visión por computadora e identificar automáticamente información dietética relevante como salida. Tam-

bién se evalúa la arquitectura de múltiples etapas y la de extremo a extremo. Donde la primera, arquitectura de múltiples etapas, consta de tres 3 etapas (análisis de imágenes de alimentos, estimación de volumen y derivación de nutrientes)

Metodología

La arquitectura de múltiples etapas para la Evaluación Dietética Basada en Visión (VB-DA) consta de tres fases: análisis de imágenes de alimentos, estimación de porciones y derivación de nutrientes. En el caso del rendimiento de las dos primeras etapas depende de los algoritmos de IA y de los conjuntos de datos, mientras que la última, de base de datos de composición nutricional, como se muestra en la figura B.1

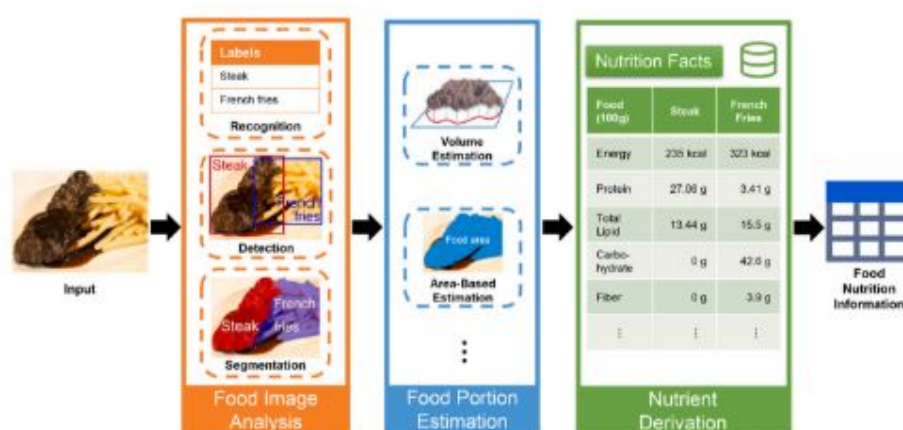


Figura 2.1: Arquitectura de múltiples etapas

El análisis de imágenes implica el reconocimiento, detección y segmentación de alimentos, donde el reconocimiento predice el tipo de alimento, la detección localiza y clasifica cada alimento mediante cuadros delimitadores, y la segmentación asigna etiquetas de alimentos a nivel de píxel, proporcionando una localización más precisa para la estimación de porciones, como muestra la figura 2.2

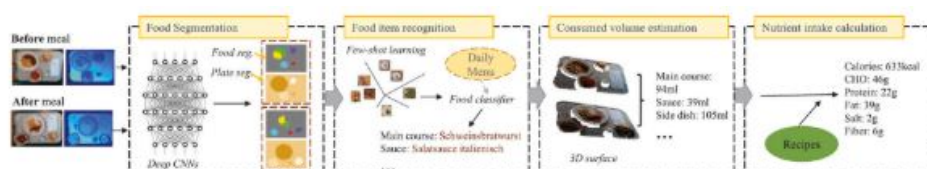


Figura 2.2: Evaluación dietética en múltiples etapas

En la arquitectura de extremo a extremo para la Evaluación Dietética Basada en Visión (VBDA) reemplaza múltiples etapas prácticas con una única red neuronal, reduciendo así la propagación y acumulación de errores, mejorando así la precisión y simplificando la optimización conjunta, evitando la necesidad de definir y optimizar etapas separadas con sus respectivas entradas y salidas. Además, de basarse en datos como el volumen de los alimentos, y adopta un marco multitarea para la estimación de nutrientes, abarcando componentes como proteínas y grasas, como se muestra en la figura 2.3

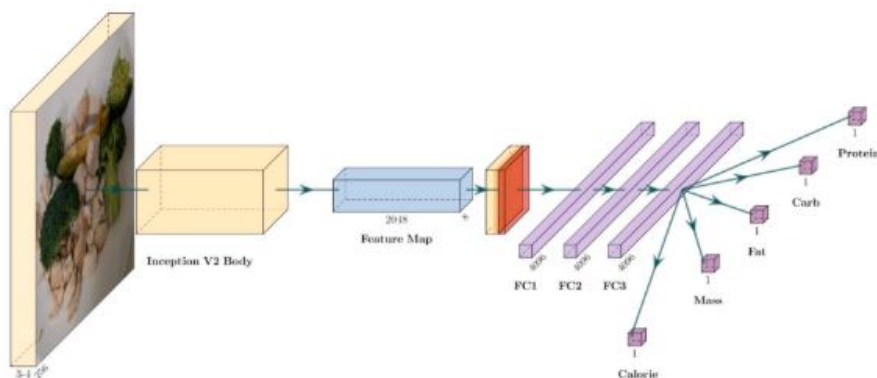


Figura 2.3: Evaluación dietética de extremo a extremo

También se usaron métodos para la detección de alimentos en la evaluación dietética implica la localización y reconocimiento de varios alimentos, utilizando técnicas avanzadas como Faster R-CNN y Mask R-CNN para mejorar la eficiencia y precisión del modelo. Estos métodos permiten reducir el área de reconocimiento y manejar propuestas regionales de manera más efectiva, facilitando la estimación de porciones y nutrientes.

Resultados y conclusiones

La precisión del análisis visual y la estimación del volumen afecta significativamente la precisión de los resultados. En la estimación de volumen, el desempeño de la porción actual de estimación del volumen aún no es satisfactorio y obtener anotaciones precisas de las porciones y la nutrición es costoso por lo que se limita la escala de los conjuntos de datos de VBDA. Aun con ello con los datos recolectados se obtuvo un MAE de 0,0933 sobre estimación de calorías, un chi-cuadrado de 0,73 en estimación de proteínas, en el caso de reconocimiento.

2.1.2. Paper 2: A Comprehensive Survey of Image-Based Food Recognition and Volume Estimation Methods for Dietary Assessment (2021)

Los autores son Tahir, G. & Loo, C. del departamento de Inteligencia Artificial, Facultad de Ciencias de la Computación y Tecnología de la Información, Universidad de Malaya, Kuala Lumpur. (Tahir & Loo, 2021)

Resumen

Problemas como la obesidad, la hipertensión y diabetes tipo 2 se deben principalmente a una mala alimentación y estilo de vida, aunque existen aplicaciones que te ayuden con ello, suelen ser complicados ya que presentan impresiones por multidimensional, la ambigüedad entre las clases por problemas de reconocimiento de los diferentes alimentos que pueden parecer similares, también están los alimentos ocultos y la calidad de cámara, dando como resultados un bajo rendimiento de los modelos para su reconocimiento. Este artículo muestra las metodologías más efectivas para el reconocimiento de alimentos y su volumen utilizando variantes de redes neuronales convolucionales (CNN) de diferentes artículos. Como el mejorar las representaciones de imágenes de alimentos se utilizan características como la forma, color, textura como métodos visuales, también reducen la complejidad eliminando características redundantes y para el volumen el proceso en su mayoría implica de dos pasos, múltiples imágenes o una sola imagen de una cámara móvil y el cálculo del volumen de alimentos a partir de una construcción 3D o un objeto de calibración. El alcance y taxonomía de este estudio se muestra el gráfico de la figura 2.4

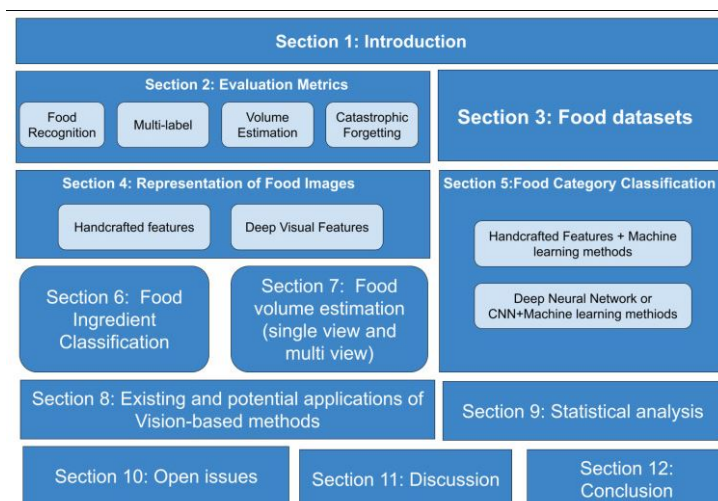


Figura 2.4: Alcance de este estudio

Metodología

Para la categorización de alimentos en su mayoría de artículos revisados se utilizaron métricas de evaluación como la matriz de confusión, Exactitud, Precisión, Recall, F1 Score. Para la representación de imágenes de alimentos se utilizaron principalmente dos métodos Handcrafted Features y Deep Visual Feature. En Handcrafted Features se extraen propiedades de las imágenes por algoritmos diseñados manualmente con el fin de capturar información relevante como la textura, forma y color de los alimentos, como lo muestra la figura 2.5. En Deep Visual Features, se usaron los siguiente clasificadores para el reconocimiento de imágenes de alimentos: máquinas de vectores de soporte (SVM), aprendizaje de núcleo múltiple (MKL) y K-vecino más cercano (KNN).



Figura 2.5: métodos de extracción: Handcrafted feature

Para la estimación del volumen de alimentos se categoriza como método de “vista de una sola imagen” y de “vista de múltiples imágenes/vídeo”, como muestra la figura 2.6

El método de vista de imagen única para estimar el volumen de alimentos requiere solo de una imagen, por lo que son más fáciles de usar, pero son menos precisos. En cambio el de vista de múltiples imágenes, es más preciso, pero más complicado de usar debido a las múltiples imágenes desde diferentes ángulos que usan para proporcionar mejores resultados.

Resultados y conclusiones

Las técnicas visuales profundas muestran un mayor rendimiento en la clasificación de alimentos, en el caso de la estimación de volúmenes de alimentos es posible calcular de manera

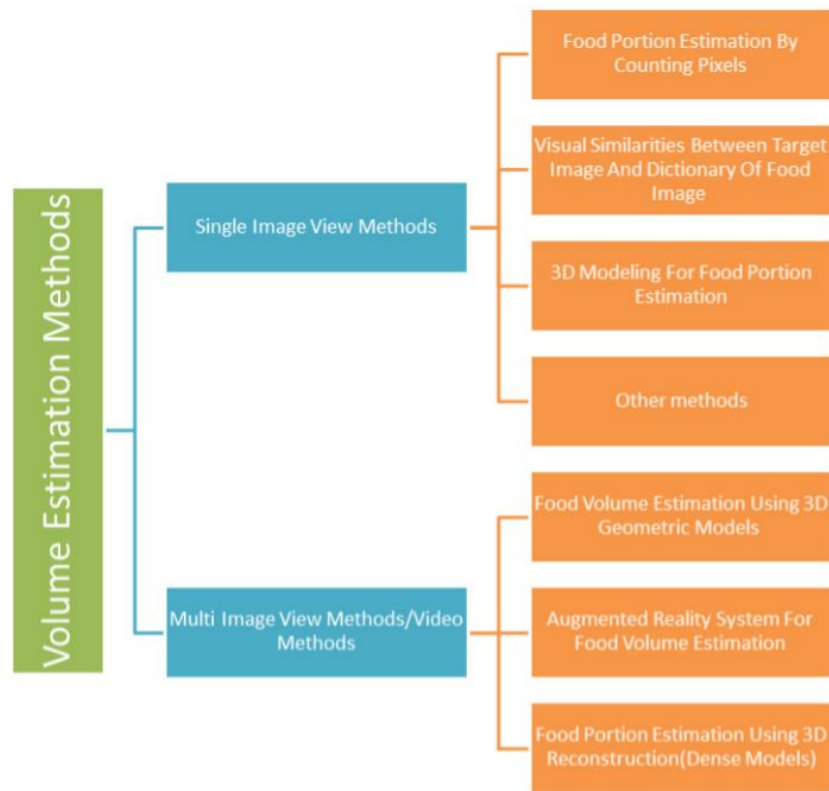


Figura 2.6: categorizacion de estimacion de volumen

precisa el tamaño de las porciones. Para el caso de reconocimiento de alimentos se usaron procesos como la adquisición, extracción, selección de características relevantes y la elección de la técnica de clasificación, como en la figura 2.7

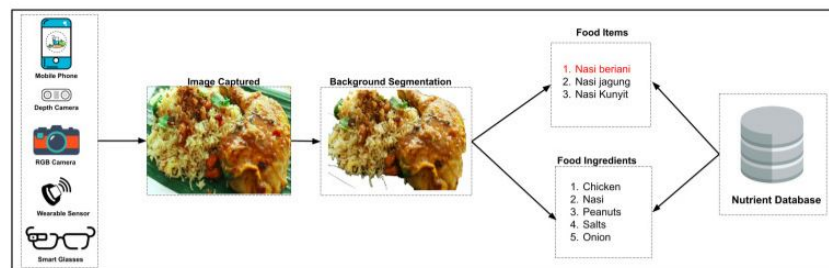


Figura 2.7: Flujo del sistema

Los artículos muestran que el 38.1 % de los conjuntos de datos son genéricos, y el 46.2 % implementan CNN para el reconocimiento de alimentos. Para estimar el volumen de alimento se recomienda usar el de vista múltiple, en la figura 2.8 se muestra la comparación de métodos de vista múltiple para la estimación del volumen de alimentos.

Reference	Year	Dataset	Results (E: Error%)	Technique
F. Zhu [141]	2010	3000 images	E: 1% 19 food items (97.2%)	"Camera calibration step and a 3D volume reconstruction step"
Xu Chang [141]	2013	14 to 20 images for multi-view method	E: 7.4% to 57.3%	Multi-view volume estimation using "Shape from Silhouettes"
Kong, Fanyu [12]	2015	6 food items	84–91%	to estimate the food portion size
Trevno, Roberto [142]	2015	120 students (n = 120 meals; 57 breakfast + 63 lunch)	74% (reliability)	Multi-View RGB images for 3D reconstruction to estimate the volume
Jia, W. Y [143]	2014	100 food samples	E: -2.80% 30%	Digital Food Imaging Analysis (DFIA)
Xu, C [36]	2013		E: 10%	ebutton is used for taking pictures, and then portion size is calculated semi-automatically by using computer software
Rhyner, D [144]	2016	6 meals	85.10%	3D MODELLING AND POSE ESTIMATION
T. Stutz [60]	2014	Rice, blinded servings	E: <33%	Multi-View RGB images, reference card and 3D model for volume estimation
Makhous et al. [145]	2020	8 food items tested	40% improvement in the accuracy of volume estimation as compared to manual calculation.	Mobile Augmented Reality System
Yuan et al. [146]	2021	Test dataset: 6 food items	E: 0.83 5.23%	Employs a mobile Structured Light System (SLS) to measure the food volume and portion size of a dietary intake.
Lo, F.P.W et al. [140]	2019	Test dataset: 11 food items	E: 15.32%	3D reconstruction from multi-view RGB images.
				3D point cloud completion from RGB and depth images.

Figura 2.8: Comparación de métodos multivista para la estimación del volumen

2.1.3. Paper 3: DeepFood: Food Image Analysis and Dietary Assessment via Deep Model (2020)

Los autores son Jiang, L., Bojia Q., Liu, X., Huang, C. & Lin, K de la Facultad de Ciencias de la Computación, Universidad McGill, Montreal, Canadá, (Jiang et al., 2020)

Resumen

Los avances en las herramientas de evaluación dietética y análisis nutricional mejoran la salud alimenticia, hay maneras de comprender los hábitos alimentarios diarios, explorando patrones nutricionales y promoviendo una dieta saludable. En este estudio, se muestra un sistema de evaluación dietética y reconocimiento de alimentos por aprendizaje profundo diseñado para analizar alimentos a partir de imágenes de comidas captadas por teléfonos inteligentes, por medio de tres pasos.

Primero, se detectan las regiones candidatas en las imágenes utilizando una red de propuesta de región (RPN) derivada del modelo Faster R-CNN, segundo, estas regiones se mapean en características y se clasifican en diferentes categorías utilizando una red neuronal convolucional profunda (CNN), lo que permite su identificación en las imágenes originales. Finalmente, el sistema analiza el contenido nutricional de los alimentos reconocidos, calculando parámetros como calorías, grasas, carbohidratos y proteínas para después generar un informe de evaluación dietética. Se utilizaron dos conjuntos de datos de imágenes de alimentos conocidos (UEC-FOOD100 y UEC-FOOD256), además de introducir un nuevo conjunto de datos basado en FOOD101. Este modelo muestra una alta precisión en el reconocimiento de ali-

mentos y genera informes de evaluación, brindando a los usuarios información valiosa sobre hábitos alimentarios saludables para una mejor salud y bienestar corporal a través de elecciones dietéticas informadas.

Metodología

El sistema basado en aprendizaje profundo en la detección de alimentos y se analiza los componentes nutricionales de cada imagen de comida, figura 2.9, el cual consta de tres pasos:

- Primero se extrae las regiones de interés (ROI) aplicando la red de propuesta de región derivada del modelo Faster R-CNN, con ello se separa los alimentos del fondo y mejora la eficiencia del modelo de detección.
- Segundo, se aplica una red neuronal convolucional (CNN) y se clasifica en diferentes categorías de alimentos, también se utiliza un módulo de regresión para localizar las coordenadas de los alimentos en la imagen.
- Tercero, se utilizan herramientas de evaluación dietética para el análisis de la nutrición de los alimentos y generar un informe de salud.

Para la detección de alimentos, se extraen los alimentos del fondo de las imágenes y se aplica Faster R-CNN para detectar las regiones relacionadas con los alimentos. Para la clasificación de los alimentos, se usó la regresión del cuadro delimitador, figura 2.10, donde se ajusta el cuadro delimitador inicial para que sus coordenadas coincidan con la verdad fundamental.

Para las métricas de evaluación se usó la precisión promedio media (mAP) para evaluar el modelo de detección, seguido de dos modelos: clasificación y regresión. La clasificación determina si un objeto existe en la imagen y la regresión determina la ubicación del objeto.

Resultados y conclusiones

En la figura 2.11 se muestran los resultados en el conjunto de datos UEC-FOOD100, donde el conjunto 1 es el experimento que utilizó todas las clases de alimentos, el conjunto 2, prueba 53 clases y el conjunto 3, con 11 clases.

En la figura 2.12 se muestran los resultados en el conjunto de datos UEC-FOOD256, donde la categoría 1 es el experimento que prueba con todas las clases de alimentos, la categoría 2, con 132 clases y la categoría 3, con 21 clases.

Los resultados en FOOD20 con bbx se muestran en la figura 2.13.

Se espera mejorar la precisión de la detección, reducir el tiempo de procesamiento, y como trabajos futuros tener la predicción del peso y crear una calculadora de dieta.

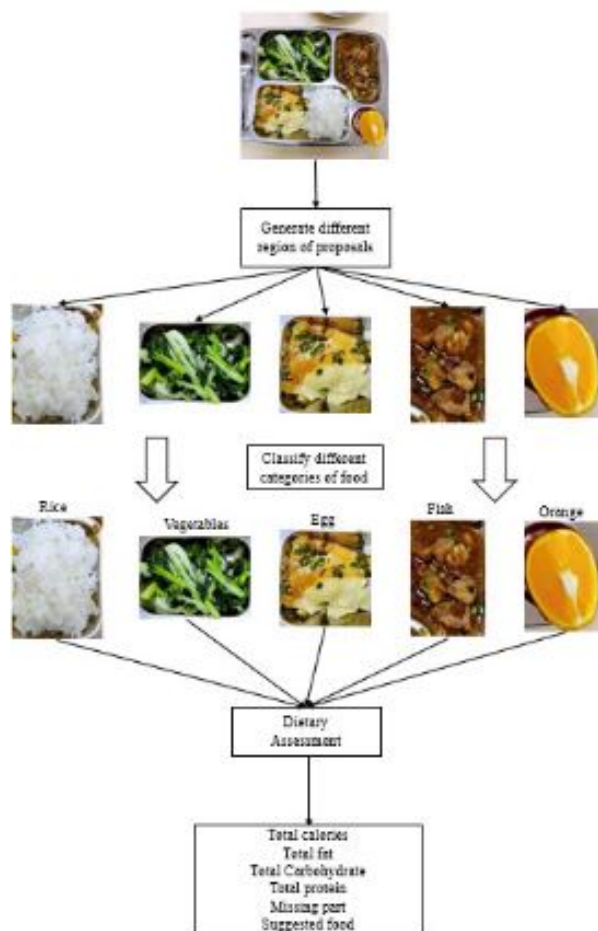


Figura 2.9: Reconocimiento de alimentos y análisis nutricional.

2.1.4. Paper 4: Image-Based Food Classification and Volume Estimation for Dietary Assessment: A Review (2020)

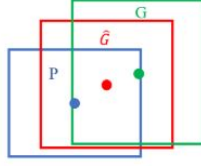
Los autores son Wen, F., Sun Y., Qiu J. & Lo, B.. del Imperial College London, Londres, Reino Unido, publicado en el año 2020. ([Lo et al., 2020](#))

Resumen

Existen varios métodos de evaluación dietética basados en imágenes. Este artículo compara modelos de reconocimiento de alimentos y la estimación de volumen/peso con respecto a la velocidad, precisión, eficiencia y limitaciones del procesamiento. También analiza la eficacia de métodos en el aprendizaje profundo para la evaluación dietética que combinan diferentes enfoques.

Metodología

Se usaron tres métodos para evaluar la ingesta dietética. El primero es el método de

**Figura 2.10:** Regresión del Bounding box

mAP(%)	Set 1	Set 2	Set 3
RCNN-based model	26.0	21.8	25.7
BP-based model	49.9	55.3	55.4
Our model	17.5	23.1	25.5

Figura 2.11: Los resultados en el conjunto de datos UEC-FOOD100

datos donde se localiza alimentos en imágenes y/o videos para reducir el almacenamiento de memoria. El segundo es el método de reconocimiento automático de alimentos, que ayuda a identificar los alimentos consumidos por los usuarios. El tercero es el método de estimación del volumen de alimentos, donde se utilizan técnicas para medir las porciones de alimento, como se muestra en la figura 2.14.

Para el reconocimiento de características de alimentos de timan en cuenta el el color, la forma y la textura, por ello hay dos enfoque principales: el reconocimiento de imágenes convencional con funciones diseñadas manualmente y el reconocimiento de imágenes de un extremo a otro con aprendizaje profundo.

En el caso del reconocimiento de imágenes convencional se usan técnicas de extracción de características como SIFT, HOG y LBP para extraer datos visuales. Para la calificación se utilizan técnicas como SVM lineal, ANN y Random Forests. Los sistemas en tiempo real utilizan técnicas de codificación como Fisher Vector para acelerar el reconocimiento. Para el aprendizaje profundo se usa las redes neuronales convolucionales (CNN) y para la estimación de volumen de alimentos de usaron metodos como: Stereo-based Approach, Model-based, Depth Camera-based, Perspective Transformation y Deep learning, como se muestra en la figura 2.15

Resultados y conclusiones

La figura 2.16 muestra la implementación del stereo-based volume estimation.

Los 12 mejores enfoques de aprendizaje profundo sobre clasificación de alimentos se muestran en la figura 2.17.

mAP(%)	Set 1	Set 2	Set 3
Our model	10.5	13.3	18.3

Figura 2.12: Los resultados en el conjunto de datos UEC-FOOD256

# of iteration	Top-1 accuracy (%)	Top-5 accuracy (%)
100	28.8	61.1
250	43.6	81.2
500	71.7	93.1

Figura 2.13: Los resultados en el conjunto de datos FOOD20

En el reconocimiento de alimentos mediante imágenes se ha demostrado que las redes neuronales profundas superan a los enfoques tradicionales basados en características manuales. Sin embargo, persisten desafíos en la estimación de la ingesta nutricional debido al tamaño de las porciones y requieren de múltiples capturas de imágenes desde diferentes ángulos.

2.1.5. Paper 5: Deep Neural Networks for Image-Based Dietary Assessment (2021)

Koroušić, B. & Mezgec, S. del Departamento de Sistemas Informáticos Ljubljana, Eslovenia, ([Mezgec & Seljak, 2021](#)).

Resumen

Con la tecnología, hoy en día, se puede registrar la ingesta con solo tomar una foto desde la cámara celular. Este artículo habla sobre el reconocimiento de alimentos en imágenes por medio de las redes neuronales profundas, utilizando la arquitectura de Nutrinet la cual es una modificación de la arquitectura AlexNet14. Aunque Nutrinet se limita a una salida de imagen de alimento, esta utiliza segmentación de imágenes ya que identifica la cantidad de alimentos y bebidas en una imagen. Los principales métodos de segmentación en este artículo son las redes totalmente convolucionales (FCN) y el otro en redes residuales profundas (ResNet). Anteriormente se ha usado extractores de características definidas manualmente pero solo han tenido precisión entre el 10% - 40% considerándose bajo por su apariencia variable de cada presentación de comida. Las redes neuronales convolucionales profundas (DCNN) son las más populares para el reconocimiento de imágenes de alimentos. Las capas convolucionales contienen filtros que aprenden y que responden a ciertas características de los datos de entrada,

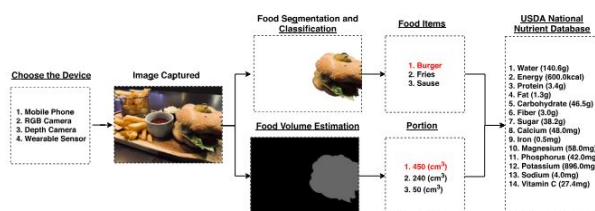


Figura 2.14: Diagrama del sistema para la evaluación dietética

Volume Estimation					
Method	Stereo-based	Model-based	Perspective transformation	Depth camera based	Deep learning
Preparation	Camera calibration (Intrinsic matrix)	Model library construction (Pre-build 3D shape models)	Strong constraints required (Camera viewing angle)	Depth camera calibration (Intrinsic matrix)	Model training
Procedure	1. Camera calibration (Extrinsic matrix) 2. 3D model reconstruction 3. 3D meshing and scale determination 4. Volume estimation	1. Food recognition 2. Model selection from the library 3. Model registration by rotating and scaling 4. Volume determination by pre-build models	1. Surface plane equation fitting 2. Perspective transformation 3. Scale determination 4. Volume calculation by geometric relationships	1. Depth map construction 2. 3D model reconstruction based on point cloud or voxel representation 3. Volume estimation	1. Single RGB image captured 2. Depth map estimation/ 3D Shape completion 3. 3D model reconstruction 4. Volume estimation

Figura 2.15: Modelos para la estimación de volumen

mientras que las capas completamente conectadas componen datos de salida de otras capas para obtener conocimientos de nivel superior a partir de ellos.

Metodología

Se utilizó la arquitectura NutriNet,[2.18](#). Para la segmentación de imágenes se tomaron dos arquitecturas: las variaciones de las redes totalmente convolucionales (FCN-32, FCN-16 y FCN-8s15) y las redes residuales profundas (ResNet), como principales tomo AlexNet14, GoogLeNet y ResNet. Para el procesamiento de datos se tomó el Food Recognition Challenge y el método de Hybrid Task Cascade con ResNet-10116, también se utilizó otras redes troncales, pero se tomó la ResNet-101 como la más adecuada. Entre los parámetros de entrenamiento para tener valores óptimos se eligió el solucionador, que determina la precisión en la clasificación; la tasa de aprendizaje, donde se define la velocidad con la cambia el parámetro de la red neuronal.

Resultados y conclusiones

Los modelos AlexNet se entrenaron utilizando un tamaño de lote de 256 imágenes y una tasa de aprendizaje base de 0,02; NutriNet utilizó un tamaño de lote de 128 imágenes y una tasa de 0,01; GoogLeNet 64 imágenes y una tasa de 0,005; y ResNet 16 imágenes y una tasa de 0,00125, figura [2.19](#)

Para entrenar el modelo de segmentación de imágenes de alimentos falsos del FCN-8, utilizamos Adam55. La precisión final del modelo FCN-8 entrenado fue del 92,18%. la figura [2.20](#) muestra, imágenes de alimentos falsos (una de cada uno de los subconjuntos de

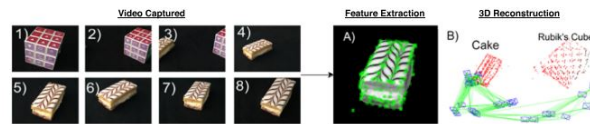


Figura 2.16: implementación de la stereo-based volume estimation

Authors	Methods	Dataset	Top 1
Kagaya et al. [27], 2014	CNN	Own dataset	73.7%
Christ. et al. [28], 2015	Patch-wise model+CNN	Own dataset	84.9%
Meyers et al. [29], 2015	CNN+inception	Food-101	79.0%
Liu et al. [30], 2016	CNN+inception	Food-101	77.4%
Pandey et al. [31], 2017	Ensemble Net	Food-101	72.1%
Liu et al. [17], 2018	CNN+inception	Food-101	77.0%
Yu et al. [32], 2018	DLA	Food-101	90.0%
Foresti et al. [33], 2018	WiSeR	Food-101	90.3%
Cui et al. [34], 2018	DSTL	Food-101	90.4%
Qiu et al. [35], 2019	PAR-Net	Food-101	90.4%
Min et al. [36], 2019	IG-CMAN	Food-101	90.4%
Tan et al. [37], 2019	EfficientNet	Food-101	93.0%

Figura 2.17: Enfoques de aprendizaje profundo, TOP 12

entrenamiento, validación y prueba), junto con las correspondientes etiquetas de predicción del modelo y de verdad sobre el terreno.

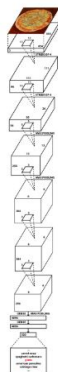


Figura 2.18: Arquitectura de red neuronal profunda NutriNet

Team Name	Placement	Average Precision	Average Recall
rssfete	1	63.4%	88.6%
simon_mezgec	2	59.2%	82.1%
arimboux	3	53.9%	73.5%
latentvec	4	48.7%	71.1%

Figura 2.19: Reconocimiento de Alimentos



Figura 2.20: Imágenes del conjunto de datos de imágenes de alimentos falsos

Anexos

Anexos A

Anexo I: Árbol del problema

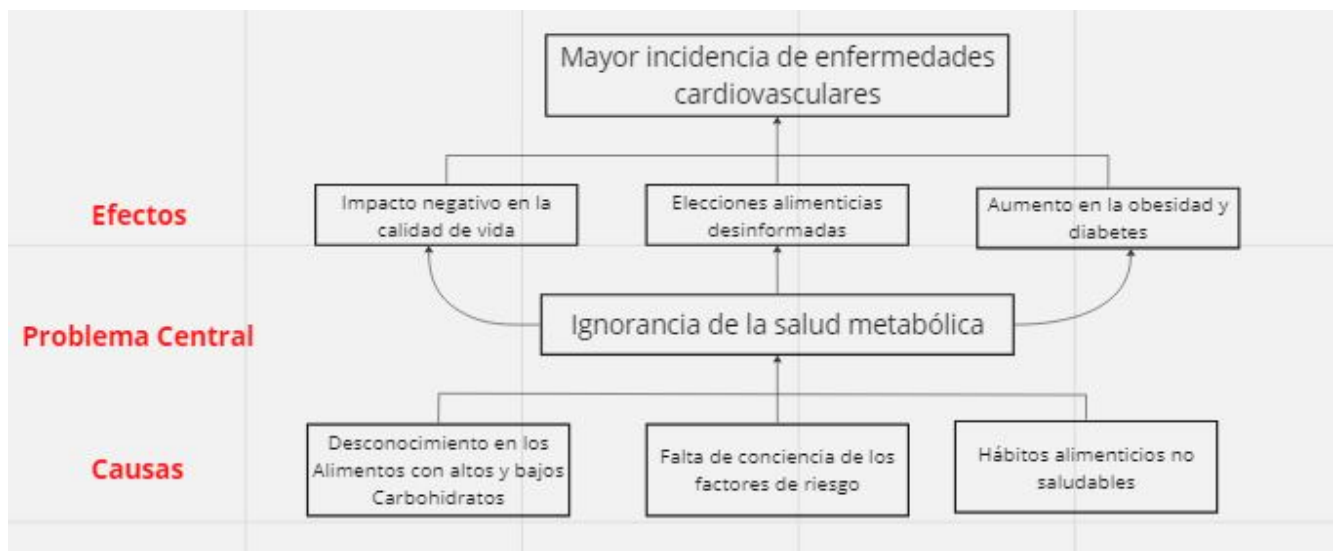


Figura A.1: Prueba de Figura

Anexos B

Anexo II: Árbol del objetivos

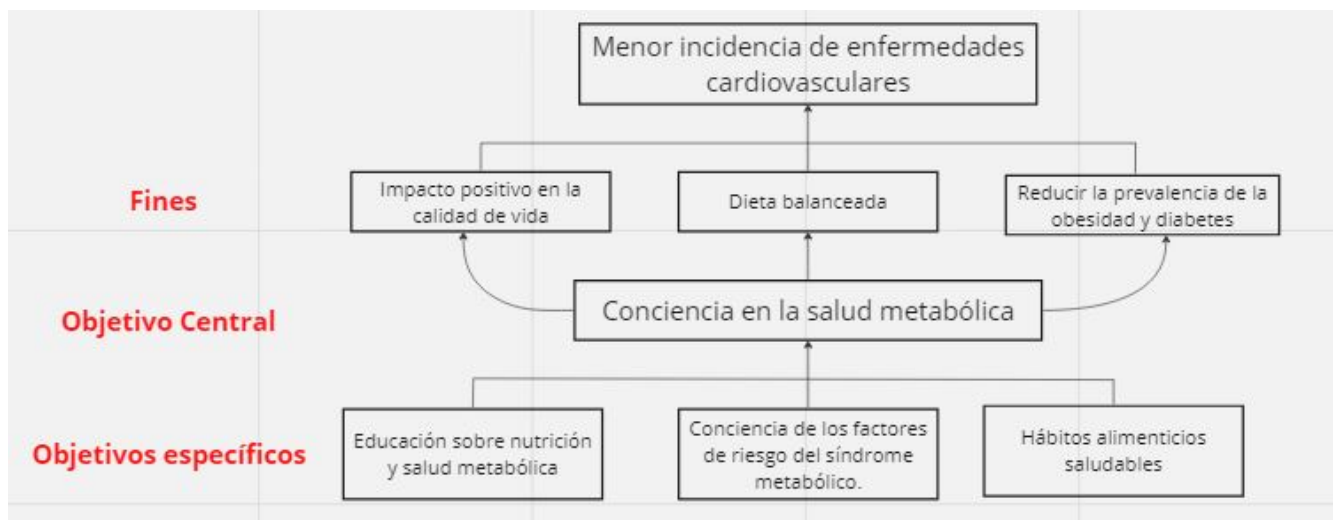


Figura B.1: Prueba de Figura

Anexos C

Anexo III: Matriz de Consistencia

PROBLEMAS	OBJETIVOS	HIPÓTESIS
Problema General	Objetivo General	Hipotesis General
¿Es posible crear una aplicación para monitorear carbohidratos aplicando vision por computadora para prevenir el síndrome metabólico en estudiantes universitarios?	Desarrollar una aplicacion que monitoree carbohidratos para ayudar a prevenir enfermedades asociadas con el síndrome metabólico en estudiantes universitarios aplicando vision por computadora.	Crear una aplicación que monitoree carbohidratos mediante visión por computadora ayuda a prevenir los riesgos asociados con el síndrome metabólico en los estudiantes universitarios
Problemas Específicos	Objetivos Específicos	Hipotesis Específicas
¿Existirá alguna base de datos de imágenes de alimentos etiquetados con su contenido de carbohidratos para poder entrenar un modelo que ayude a prevenir el síndrome metabólico?	Encontrar y/o construir una base de datos de imágenes de alimentos etiquetados con su contenido de carbohidratos para el entrenamiento de un modelo que ayude a prevenir el síndrome metabólico.	Una base de datos de imágenes de alimentos etiquetados con su contenido de carbohidratos permite entrenar un modelo que ayude a prevenir el síndrome metabólico.
¿Cómo medir la efectividad de la aplicación de visión por computadora que monitorea los carbohidratos para ayudar a prevenir el síndrome metabólico en estudiantes universitarios?	Determinar la efectividad de la aplicacion que ayuda a prevenir el síndrome metabólico en estudiantes universitarios monitoreando los carbohidratos a traves de vision por computadora	La efectividad de la aplicacion de visión por computadora que monitorea los carbohidratos ayuda a prevenir el síndrome metabólico en estudiantes universitarios
¿Cuál es la mejor arquitectura de redes neuronales convolucionales (CNN) para poder entrenar un modelo de monitoreo de carbohidratos?	Determinar la mejor arquitectura de redes neuronales convolucionales (CNN) para entrenar un modelo de monitoreo de carbohidratos.	La mejor arquitectura de redes neuronales convolucionales (CNN) permite entrenar un modelo de monitoreo de carbohidratos
¿Cuál es el método de preprocesamiento más adecuado para mejorar la calidad de las imágenes de alimentos para el monitoreo de carbohidratos que presentan un alto nivel de ruido?	Determinar la mejor arquitectura de redes neuronales convolucionales (CNN) para entrenar un modelo de monitoreo de carbohidratos.	Mediante la aplicación de técnicas de preprocesamiento de imágenes, se logra mejorar la precisión del modelo de monitoreo de carbohidratos al reducir la interferencia del ruido en las imágenes de alimentos.

Tabla C.1: Matriz de consistencia. Fuente: Elaboración propia

BIBLIOGRAFÍA

- Albornoz Lopez, R., & Perez Rodrigo, I. (2012). Nutrición y síndrome metabólico. *Nutr. clín. diet. hosp*, 32(3), 92-97.
- CISAN. (2022). ÍNDICE GLUCÉMICO. <https://infoalimentos.org.ar/informes/documentos-tecnicos/409-indice-glucemico-documentos>
- Dobrowolski, P., Prejbisz, A., Kuryłowicz, A., Baska, A., Burchardt, P., Chlebus, K., Dzida, G., Jankowski, P., Jaroszewicz, J., Jaworski, P., et al. (2022). Metabolic syndrome—A new definition and management guidelines. *Arterial Hypertension*, 26(3), 99-121.
- Jiang, L., Qiu, B., Liu, X., Huang, C., & Lin, K. (2020). DeepFood: food image analysis and dietary assessment via deep model. *IEEE Access*, 8, 47477-47489.
- Lemieux, I., & Després, J.-P. (2020). Metabolic syndrome: past, present and future.
- Lo, F. P. W., Sun, Y., Qiu, J., & Lo, B. (2020). Image-based food classification and volume estimation for dietary assessment: A review. *IEEE journal of biomedical and health informatics*, 24(7), 1926-1939.
- Mezgec, S., & Seljak, B. K. (2021). Deep neural networks for image-based dietary assessment. *JoVE (Journal of Visualized Experiments)*, (169), e61906.
- Tahir, G. A., & Loo, C. K. (2021). A comprehensive survey of image-based food recognition and volume estimation methods for dietary assessment. *Healthcare*, 9(12), 1676.
- Vera-Ponce, V. J., Talavera, J. E., Torres-Malca, J. R., La Cruz-Vargas, D., & Jhony, A. (2022). Combinaciones de síndrome metabólico y riesgo de diabetes mellitus. *Revista Cubana de Medicina Militar*, 51(1).
- Wang, W., Min, W., Li, T., Dong, X., Li, H., & Jiang, S. (2022). A review on vision-based analysis for automatic dietary assessment. *Trends in Food Science & Technology*, 122, 223-237.