

UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA

FACULTAD DE MEDICINA Y CIENCIAS BIOMÉDICAS

SECRETARÍA DE INVESTIGACIÓN Y POSGRADO



UNIVERSIDAD AUTÓNOMA DE
CHIHUAHUA

**“Detección de Factores de Riesgo de Muerte de Cuna por Medio de
Inteligencia Artificial”**

POR:

ALEXIS SALVADOR ESCOBAR BESERRA

TESIS PRESENTADA COMO REQUISITO PARA OBTENER EL GRADO DE:
**MAESTRÍA EN FORMACIÓN E INNOVACIÓN PARA PROFESIONALES DE LA
SALUD**

CHIHUAHUA, CHIH., MÉXICO

MAYO DE 2025



Universidad Autónoma de Chihuahua
Facultad de Medicina y Ciencias Biomédicas
Secretaría de Investigación y Posgrado.



La tesis **“Detección de Factores de Riesgo de Muerte de Cuna por Medio de Inteligencia Artificial”** que presenta Alexis Salvador Escobar Beserra, como requisito parcial para obtener el grado de: Maestría en Formación e Innovación para Profesionales de la Salud, ha sido revisada y aprobada por la Facultad de Medicina y Ciencias Biomédicas.

DR. SAID ALEJANDRO DE LA CRUZ REY
Secretario de Investigación y Posgrado
Facultad de Medicina y Ciencias Biomédicas
Universidad Autónoma de Chihuahua

DRA. HAYDEE PARRA ACOSTA
Profesora Coordinadora Académica
Facultad de Medicina y Ciencias Biomédicas. UACH.

DRA. KARINA RODRÍGUEZ CARMONA
Directora de Tesis
Facultad de Medicina y Ciencias Biomédicas. UACH.

DR. JOSÉ LUIS HERRERA AGUILAR
Co- director de Tesis
Facultad de Ingeniería. UACH.

DR. ROGELIO RODRÍGUEZ MAESE
Asesor
Centro de Investigación en Materiales Avanzados

DR. CARLOS ROBERTO CERVANTES SÁNCHEZ
Asesor
Facultad de Medicina y Ciencias Biomédicas. UACH.

DR. CARLOS ARZATE QUINTANA
Asesor
Facultad de Medicina y Ciencias Biomédicas. UACH.

Karina Rodríguez C.

Se certifica, bajo protesta de decir verdad, que las firmas consignadas al pie del presente documento son de carácter original y auténtico, correspondiendo de manera inequívoca a los responsables de las labores de dirección, seguimiento, asesoría y evaluación, en estricta conformidad con lo dispuesto en la normativa vigente de esta institución universitaria.

Resumen

Esta investigación se enfoca en desarrollar dos programas basados en Inteligencia Artificial (IA) para detectar factores de riesgo de muerte de cuna. Se buscó obtener un *accuracy* mayor a 85% en ambos modelos en sus respectivas fases de prueba utilizando imágenes de un maniquí y de un bebé real. El primer programa se enfoca en la detección de posiciones usando cinco categorías: Normal (boca arriba o decúbito dorsal), Lateral, Boca Abajo, Tapado hasta el abdomen y ASSB (detección de objetos y cabeza tapada). Se utilizó el modelo ResNet y una base de datos de 4,000 imágenes; el segundo programa se enfocó en la detección de puntos rojos en el abdomen con el modelo YOLOv8 utilizando 500 imágenes. Dicha detección tiene el fin de medir las respiraciones en el abdomen. En la fase de prueba del programa de detección de posiciones se obtuvo un *accuracy* de 99.4% con el maniquí y 73% con el bebé, mientras que el modelo de detección de puntos rojos alcanzó un *accuracy* de 97%. El resultado en ambos programas fue satisfactorio, sin embargo, en base a los resultados obtenidos con las imágenes del bebé del primer programa, se ve la necesidad de crear una base de datos de imágenes de diversos bebés para un mejor entrenamiento en entornos reales. Ambos modelos tienen la finalidad de ser implementados en un futuro en un sistema integral para el monitoreo no invasivo en tiempo real.

Palabras clave: Diseño, Programación, ResNet, YOLOv8.

Abstract

This research focuses on the development of two Artificial Intelligence (AI) programs to detect risk factors for Sudden Infant Death Syndrome (SIDS). The objective was to achieve an *accuracy* greater than 85% for both models in their respective testing phases, which utilized images of both a mannequin and a real infant. The first program is dedicated to position detection, classifying five categories: Normal (supine), Lateral, Boca Abajo (Prone), Tapado (Covered to the abdomen), and ASSB, which involves the detection of objects and a covered head. This model was developed using ResNet and a dataset of 4,000 images. The second program, built with YOLOv8 and a 500-image dataset, focuses on detecting red dots on the abdomen to measure respiration. During the testing phase, the position detection program obtained 99.4% *accuracy* with the mannequin and 73% with the infant, whereas the red dot detection model reached 97% *accuracy*. While the overall results are satisfactory, the performance with the real infant images highlights the necessity of creating a larger, more diverse database of infants to improve training for real-world environments. The ultimate goal for both models is their future implementation into an integrated, non-invasive, real-time monitoring system.

Keywords: Design, Programming, ResNet, YOLOv8.



Chihuahua, Chih., a 18 de marzo de 2025

Dr. Said Alejandro de la Cruz Rey
Secretario de Investigación y Posgrado
Facultad de Medicina y Ciencias Biomédicas
P R E S E N T E:

Por medio de la presente nos permitimos informar a usted que el C. Alexis Salvador Escobar Beserra, con número de matrícula 290208, ha concluido la elaboración de la tesis **“Detección de Factores de Riesgo de Muerte de Cuna por Medio de Inteligencia Artificial”**, como requisito para obtener el grado de: **Maestría en Formación e Innovación para Profesionales de la Salud.**

Así mismo, manifestamos que la tesis ha sido revisada y aprobada por los abajo firmantes, miembros del Comité de Grado.

Sin otro particular, quedamos de usted.

ATENTAMENTE

“MENTI DA LUCEM, MANIBUS ARTEM”

Karina Rodriguez C.

Dra. Karina Rodríguez Carmona
Directora de tesis. FMyCB

[Signature]

Dr. José Luis Herrera
Co-director de tesis. FI

[Signature]

Dr. Rogelio Rodríguez Maese
Asesor. CIMAV

[Signature]

Dr. Carlos Roberto Cervantes Sánchez
Asesor. FMyCB

[Signature]

Dr. Carlos Arzate Quintana
Asesor. FMyCB

Esta investigación se desarrolló en la Facultad de Medicina en la Universidad Autónoma de Chihuahua en ciudad Chihuahua, Chihuahua.

Registro de comité de investigación (SIP): CI – 030 - 23

Esta maestría fue realizada gracias al apoyo de beca CONACYT.

Dedicatoria

Dedico este trabajo a los que me apoyaron en cada caída, a los que me acompañaron en cada victoria, y a la misma vida, por recordarme que todo esfuerzo vale la pena.

Agradecimientos

Agradezco primeramente a Dios por la oportunidad de vivir esta experiencia única, a mi familia, por su amor incondicional y por apoyar mis decisiones, a mis asesores, quienes con su guía y mucha paciencia me apoyaron en cada etapa del proceso, a mis amigos, por su apoyo constante y palabras de ánimo; de manera especial, agradezco a la familia que me permitió incluir a su bebé en esta investigación demostrando su generosidad y confianza invaluable.

Y, finalmente, a mí, por demostrar que soy capaz de lograr las metas que me propongo.



Índice

INTRODUCCIÓN	1
MARCO TEÓRICO.....	2
<i>Muerte de Cuna</i>	2
<i>Factores de Riesgo de SUID</i>	3
<i>Programación</i>	4
ANTECEDENTES	9
Avances Tecnológicos	9
Uso de Sensores	9
Uso de Ultrasonido	9
Uso de Infrarrojo	10
Uso de Cámaras e IA	10
DELIMITACIÓN DEL OBJETO DE ESTUDIO	12
<i>Planteamiento del Problema y Pregunta de Investigación</i>	12
<i>Hipótesis Predictoras</i>	13
<i>Objetivos</i>	13
Objetivo General.....	13
Objetivos Específicos	14
JUSTIFICACIÓN	14
MATERIALES Y MÉTODOS	15
<i>Diseño del Estudio</i>	15
<i>Unidad de Estudio</i>	15
Tamaño de Muestra	15
Materiales	16
Definición de Variables	16
<i>Técnicas y Procedimientos</i>	17
Etapa 1. Diseño de Set-up:	18
Etapa 2: Creación de Base de Datos	19
Etapa 3: Diseño de Instrumento.....	21
Etapa 4: Validez del Instrumento	21
Etapa 5: Creación de Programa de IA para la Detección de Factores de Riesgo	24



Etapa 6: Implementación de Programación con Fotos de Bebé (Fase de Prueba)	25
Etapa 7: Creación de Programa para la Detección de Respiraciones	27
<i>Consideraciones Éticas</i>	29
RESULTADOS	30
<i>Introducción</i>	30
<i>Análisis de Instrumento</i>	30
Validación de Contenido	30
Validación de Confiabilidad	33
<i>Caracterización de Muestra</i>	34
<i>Análisis de Programa para la Detección de Posiciones</i>	37
Primer Nivel de Análisis: Imágenes de Maniquí	37
Segundo Nivel de Análisis: Imágenes de Maniquí	39
Segundo Nivel de Análisis: Correlación con Maniquí	40
Tercer Nivel de Análisis: Imágenes de Maniquí	41
Segundo Nivel de Análisis: Imágenes de Bebé	42
Tercer Nivel de Análisis: Imágenes de Bebé	44
<i>Análisis de Programa para la Detección de Respiraciones</i>	44
Primer Nivel de Análisis	45
Segundo Nivel de Análisis	46
Tercer Nivel de Análisis	48
<i>Tesis</i>	49
DISCUSIÓN	50
CONCLUSIONES	53
BIBLIOGRAFÍA	53
ANEXOS	57
<i>Anexo 1. Formato de carta de consentimiento</i>	57
<i>Anexo 2. Aprobación de Protocolo de Investigación</i>	59



INTRODUCCIÓN

La muerte de cuna (Crib Death en inglés) o la Muerte Súbita e Inesperada del Lactante (SUID por sus siglas en inglés), es un término amplio utilizado para describir cualquier muerte repentina, ya sea explicada o no, que ocurre durante el primer año de vida. Este término abarca Síndrome de Muerte Súbita del Lactante (SMSL o SIDS en inglés), Asfixia Accidental en un Ambiente para Dormir (ASSB por sus siglas en inglés) y otras muertes por causas desconocidas [1,2].

La CIE-10 (Clasificación Estadística Internacional de Enfermedades y Problemas Relacionados con la salud, Decima revisión) lo clasifica bajo el código R95 dentro de “Causas de mortalidad mal definidas y desconocidas” con el nombre de Síndrome de la muerte súbita infantil [3].

El SMSL es la muerte abrupta e inexplicable de un bebé menor de un año de edad al que no se encuentra razón de su fallecimiento [4,5]. Con el paso de los años se han buscado soluciones para reducir la incidencia de muerte de cuna; sin embargo, aún no se cuenta con un equipo para uso doméstico de bajo costo, que sea confiable y, sobre todo, que no tenga efectos negativos como radiaciones o afecciones a la piel.

Actualmente se han desarrollado diversos dispositivos que ayudan en la detección de muerte de cuna entre los que destacan el uso de sensores Lilypads y acelerómetros debido principalmente a lo accesibles que son económicamente [6].

Otros equipos implementan la detección de posiciones o movimientos mediante sensores infrarrojos [7,8], ultrasonido [9] y cámaras que emplean Inteligencia Artificial (IA) [10–14], sin embargo, la mayoría de estos métodos aún se encuentran en fase experimental y muy pocos están disponibles para uso doméstico, además los equipos más modernos suelen ser de alto costo y no se comercializan en México.



En este trabajo se presenta una alternativa innovadora, basada en el desarrollo de 2 programas implementando IA: el primero detecta posiciones de riesgo y objetos que propician una muerte de cuna y muerte por asfixia, y un segundo programa realizaría la detección de 2 puntos rojos para la detección de respiraciones. Ambos programas fueron entrenados con redes neuronales convolucionales (CNNs) utilizando bases de datos específicas incluyendo imágenes de maniquí y de un bebé real, esto con el fin de evaluar su eficacia tanto en entornos controlados como en situaciones reales.

Esta investigación es la base para el desarrollo futuro de un sistema no invasivo implementando una computadora y una cámara, siendo seguros para un bebé. El objetivo es que los programas diseñados logren superar la precisión de 85% y obtener modelos finales para ser implementados en un futuro en un sistema integral para reducir la tasa de incidencia de muerte de cuna. Cabe mencionar que la metodología y el diseño de set-up fueron propios y en esta investigación los bebés hacen referencia a niños desde recién nacidos hasta un año de edad.

MARCO TEÓRICO

Muerte de Cuna

En Estados Unidos (EU) existe una reducción de SUID gracias a los avances tecnológicos y la capacitación a las familias de 154.6 por cada 100,000 nacidos vivos en 1990 a 90.1 en 2019 [2], [10]. Actualmente en EU se calcula que al año mueren cerca de 3,500 infantes por SUID [1], de los cuales 1,398 son por SMSL [4] y aproximadamente 980 (28%) de las muertes son causadas sofocación accidental [10], sin embargo en México no se cuenta con un registro oficial de la incidencia de SUID debido a la falta de programas que se dediquen a la muerte de cuna como EU.

Un estudio realizado para estimar la incidencia de SUID en México entre 2005 a 2020 calculó que en 2020 la tasa era de 29 por cada 100,000 nacidos vivos; la tasa de SMSL fue de 26.7 por cada 100,000 y la de ASSB 0.1 por cada 100,000 [2].



La tasa de incidencia crece en los países en vías de desarrollo debido a que no cuentan con la misma información ni tecnología que los países desarrollados según la Organización Mundial de la Salud [15].

Factores de Riesgo de SUID

Hablando de los principales factores de riesgo se estima que la mayor tasa de incidencia de SUID (aproximadamente el 90%) ocurren dentro de los primeros seis meses de vida, teniendo un alza entre los dos y cuatro meses [4,16]. Basándonos en esta información, se buscó un bebé que tuviera entre 2 y 4 meses de edad para la fase de prueba. Dado que la fase de sueño más prolongada es durante los primeros seis meses de vida (14.5 horas de sueño en promedio al día) es muy importante mantener una vigilancia constante al bebé [17]. Además, se ha observado que los bebés varones tienen un mayor riesgo con una proporción de 2:1 en comparación con las niñas [16].

Otro factor que consideran varios autores son las posiciones al dormir, ya que algunos determinan que la posición boca abajo (decúbito prono) triplicaba la incidencia de SUID en el bebé en comparación con la posición boca arriba (decúbito supino o dorsal) [16]; también se señala que la posición de lado es factor de SMSL [4]. Gracias a esta información es que se considera la detección de las posiciones boca arriba, de lado y boca abajo en esta investigación.

Dormir con los padres se considera otro factor de riesgo cuando se comparte cama y/o no existe una circulación de aire, ya que el aire exhalado por los adultos contiene mayores concentraciones de dióxido de carbono (CO_2), esto incrementa el riesgo de hipoxia en el bebé al inhalar aire con alto contenido de CO_2 y bajo nivel de oxígeno.

Existen muchas recomendaciones para prevenir SMSL, algunas de ellas son: dormir en la habitación de los padres en una cuna con circulación de aire, llevar a cabo la lactancia en la medida de lo posible, uso de chupete al dormir, que la base de la cama no esté inclinada más de 10 grados, promover pláticas informativas sobre la muerte de cuna y cómo prevenirla, usar una base para dormir rígida, utilizar prendas en vez de mantas, ya que estas últimas pueden provocar alguna



asfixia si se llevan a la cara, entre otros [1]. Otros factores extrínsecos que aumentan el riesgo de SUID son las temperaturas ambientes bajas, temperaturas muy elevadas y la exposición al humo de tabaco [4].

El modelo de triple riesgo es una teoría ampliamente aceptada que da explicación a una muerte de cuna, esta teoría menciona que ocurre cuando tres factores coinciden simultáneamente. El primero es la vulnerabilidad del bebé, que es una cuestión intrínseca que incluye condiciones subsecuentes como disfunción del control cardiorrespiratorio o problemas neurológicos; el segundo es el periodo crítico del desarrollo, generalmente los primeros meses (principalmente entre los meses dos y cuatro de vida) y es por ello que para la fase de prueba se procurará obtener imágenes de un bebé de esa edad. Finalmente los factores externos como los que se mencionaron en este apartado [1].

La ASSB suele ser causada principalmente por descuidos de los cuidadores como dejar objetos en la cuna (juguetes, cobijas, sábanas), colocar ropa no adecuada al bebé e incluso cubriéndolo con mantas o cobijas cerca de su cara [16]. Es por esta razón que nos enfocaremos también en la detección de cobijas que se encuentren de manera superior al tronco, ya que es un riesgo de asfixia.

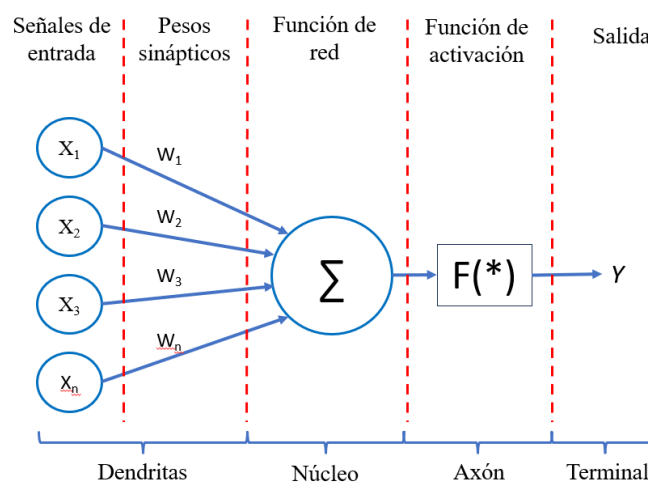
Programación

La IA es la capacidad de una máquina de realizar tareas que hacemos los humanos, son diversos procesos matemáticos que crean una forma de inteligencia [18,19]. Para entender lo que es la IA primero debemos hablar de lo que es una red neuronal: una red neuronal tiene funciones como procesamiento de información en paralelo, memoria distribuida y adaptabilidad del entorno [20]. La Figura 1 muestra un modelo de red neuronal con sus principales partes y similitud con una neurona, las principales partes que componen una red neuronal son:

- Señal de entrada: son los valores que serán evaluados en base.
- Pesos sinápticos: valores numéricos que determinan la importancia o influencia que tiene cada entrada en la salida de una neurona.

- Función de red: es la encargada de realizar la suma de las señales de entrada y realizar las operaciones propuestas en el modelo.
- Función de activación: limita la salida de información de la neurona, determina cuales son las categorías o ítems de salida.
- Capa de salida: señala la identidad de la información proporcionada.

Figura 1: Modelo genérico sobre componentes de una red neuronal.



Nota: La red neuronal consta principalmente de cinco etapas parecidas a una neurona (referencia al nombre de redes neuronales) en donde se muestra el ingreso de diversas señales de entrada, una fase de procesamiento, una fase que limita la salida de información y una salida (Fuente propia).

La mayoría de los lenguajes de programación son en inglés, por lo que siempre en ingeniería la terminología se utiliza en ese idioma. En esta investigación tanto el lenguaje de programación como los conceptos utilizados son en inglés. Para los proyectos de tecnología y electrónica, “set up” se refiere al proceso de preparar y organizar todos los componentes para que un sistema funcione correctamente, esto incluye tanto hardware como software.

El aprendizaje automático (Machine learning) es un modelo de aprendizaje que aprende modelos matemáticos a partir de datos de muestra (fase de entrenamiento) para crear predicciones sobre datos aun no vistos (fase de validación y prueba), principalmente se evalúan 2 métricas para



evaluar su rendimiento: el tiempo que tarda en dar un resultado y el *accuracy*. La primera determina que tan bueno es el algoritmo para ser aplicados a gran escala y esto depende tanto del tamaño como del peso de la red neuronal, así como de las características computacionales.

Accuracy es un término en inglés utilizado en redes neuronales que mide la proporción de predicciones correctas (verdaderos positivos o negativos) sobre el total de predicciones, determina la utilidad y aplicabilidad en el mundo real [21]. El termino *Loss* (pérdida) indica que tan bien o mal está desempeñándose el modelo, es una medida cuantitativa del error entre las predicciones del modelo y los valores reales que utiliza el programa para ajustar pesos y así mejorar los modelos.

Las Redes Neuronales Convolucionales (CNN por sus siglas en inglés) es un método de machine learning que sirve para el procesamiento de imágenes, tratando de imitar al córtex visual del cerebro humano (19,20). Este tipo de modelo es muy implementado gracias a su gran capacidad para aprender patrones complejos en imágenes, mejorando la precisión de detección de objetos, así como clasificación y segmentación de imágenes [12].

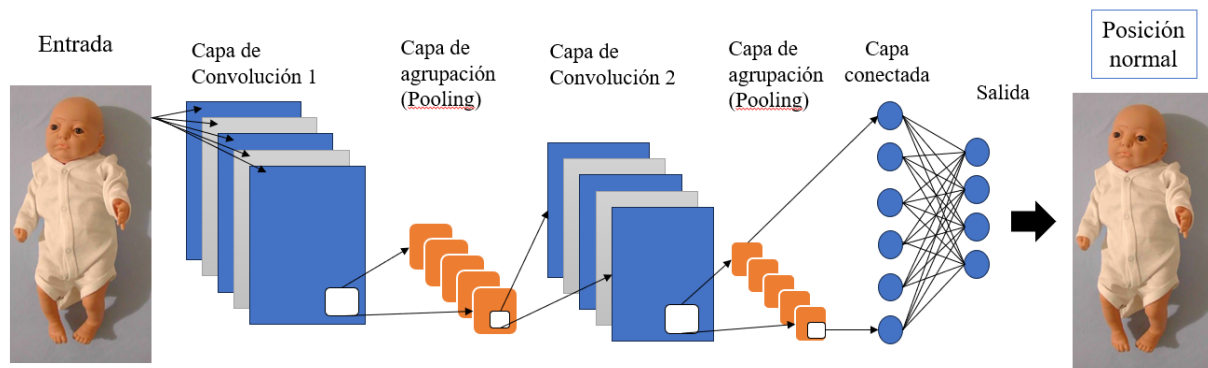
Este tipo de programaciones pueden llegar a ser de diversos números y tipos de capas, así como diferentes conexiones entre las mismas debido a que cada proyecto es diferente y no existe un patrón establecido. Las primeras capas detectan propiedades y formas básicas que se van especializando; después tenemos las capas profundas, las cuales detectan formas complejas como rostros, siluetas o formas en específico [20].

Las capas de entrada son las primeras, reciben la imagen en forma de píxeles y luego las pasan a las capas de convolución, mismas que aplican filtros en pequeñas regiones de la imagen para crear y resaltar mapas de características específicas. Las capas pooling reducen la dimensión de los mapas de características, conservando las más relevantes y mejorando la lectura.

Después tenemos las capas profundas, la primera es la capa conectada que combina las características extraídas en las capas anteriores para crear una representación de alto nivel, y por último la capa de salida que genera la predicción final, asignando una imagen a una categoría en

específico [12]. La Figura 2 muestra un diagrama simbólico de cómo se lleva a cabo este tipo de red neuronal utilizando una imagen de esta investigación.

Figura 2: Diagrama de funcionalidad de la CNN.



Nota: Ejemplo de entrenamiento de IA por medio de CNN. Primero entran las imágenes, después pasan por las capas de convolución detectando características específicas para después pasar por una capa de pooling, que reducen los mapas de características seleccionando solo las más relevantes para mejorar la lectura (en este ejemplo existe otra capa de convolución y pooling realizando las mismas funciones de cada una y se pueden modificar el número de capas); después proceden las capas conectadas que combinan las características para crear una representación de alto nivel. Por último, una capa de salida muestra la predicción final (diseño propio).

Para la creación de la IA se utilizó el lenguaje de programación *Python*, el cual es un lenguaje de libre acceso. El batch (tamaño de lote) es el número de imágenes (muestras) que se procesan en cada iteración (una iteración se realiza cuando procesa cada una de las imágenes del batch y ajusta valores).

Después de procesar el número de muestras que se determinen, el programa realiza una actualización de sus parámetros y volverá a procesar el número de muestra una y otra vez hasta terminar el conjunto de datos total terminando una Epoch (épocas) mismo que representa una pasada del conjunto completo de datos a través de la red neuronal.

Existen diversos métodos para mejorar el aprendizaje como los son el Dropout y “batch normalization”, donde según la investigación de Garbin et al. [21] el uso de “batch normalization” ayuda significativamente a acelerar el entrenamiento de redes profundas, mientras que el Dropout ayuda a que no exista el sobre ajuste siempre y cuando el tamaño de lote sea grande (256 muestras o más). Dichas técnicas serán implementadas y evaluadas en esta investigación.

Para el diseño de las redes neuronales no existe un patrón definido para establecer el número de capas ni los tipos de conexión, por lo que cada proyecto que utiliza este tipo de red neuronal es independiente y se somete a cambios constantes para obtener el mejor resultado [20]. Actualmente se utilizan programas pre-entrenados que han demostrado ser efectivas y con una alta tasa de asertividad en el proceso de imágenes como son: *LeNet*, *AlexNet*, *VGG*, *GoogleNet*, *ResNet* y *DenseNet*.

ResNet (Residual Network) es una arquitectura de red neuronal profunda que introduce conexiones residuales. A diferencia de los otros métodos, permite entrenar redes más profundas sin problema de degradación del gradiente (pérdida de información). Esto ayuda a la capacidad de la red de aprender características complejas, lo que ayudará en el análisis de las características que se buscan en esta investigación [20]. *YOLO* es un modelo que pertenece a los detectores de una capa utilizado principalmente en la detección de objetos en tiempo real. Es altamente reconocido por la precisión y la velocidad para detectar múltiples objetos en imágenes o video y utiliza una sola etapa que gestiona una generación de cuadros delimitados en las imágenes y predice clases en una sola evaluación, por lo que es un modelo sencillo y rápido [12].

YOLOv8 es un método más actual basado en clasificación que tratan la detección de objetos como un problema de regresión, manejando simultáneamente cuadros y probabilidades lo que simplifica la arquitectura de la red y reduce el número de parámetros [14]. El modelo *YOLOv8* mide cuatro parámetros: *accuracy* que mide que tan bien el modelo evita falsos positivos, *Recall* (sensibilidad) que mide que tan bien el modelo encuentra todos los verdaderos positivos, *mAP 50* que mide la precisión cuando el solapamiento es del 50% o más, y *mAP 50-95* que mide la precisión en ese rango.



ANTECEDENTES

Avances Tecnológicos

Kaler et al. [5] menciona que en 1969 se descubrió el SMSL, creando el primer equipo para su detección. Con el paso del tiempo se han diseñado diversos equipos para prevenir SUIDs; a continuación, se explicarán algunos:

Uso de Sensores

Desde una perspectiva histórica, Alfred Steinschneider en su publicación de 1972 señalaba que la apnea era la principal causa de SMSL, lo que provocó la creación de los primeros monitores de apnea, sin embargo, su publicación fue desacreditada 25 años después debido a que no se tomaban en cuenta más variables para la detección de SMSL como la temperatura, el movimiento y la frecuencia cardíaca [22]. Con el paso de los años se han mejorado los equipos para la detección de signos vitales por lo que actualmente existen equipos que captan diversos parámetros para la prevención de SUID.

Los acelerómetros son sensores de contacto que detectan movimientos. Actualmente se utilizan los lilypads, que son placas electrónicas flexibles con diversos sensores integrados (acelerómetros, sensores de temperatura, etc.) que se pueden colocar en la prenda del bebé para la recolección de información como temperatura, saturación de oxígeno y movimientos [6]; desafortunadamente y en base a las críticas de productos comercializados en plataformas digitales, las tasas de efectividad y confiabilidad de estos dispositivos suelen ser bajas. Por esta razón, muchos investigadores implementan varios sensores para mejorar la precisión y monitorear mayor número de variables para cuidar al bebé logrando obtener mejores resultados, la desventaja es la elevación de costo y pueden llegar a irritar la piel [23].

Uso de Ultrasonido

Una innovación es la invención de equipos que detectan la respiración por medio de ondas ultrasónicas de hasta 500 MHz, estas ondas son transmitidas por un sensor de ultrasonido ubicado a cierta distancia que detecta las ondas cuando rebotan con el bebé [9]. Actualmente aún se hacen investigaciones para detectar si existe algún efecto perjudicial hacia el bebé.



Uso de Infrarrojo

Otros dispositivos son los sensores Grib-Eye (lente infrarroja) los cuales ayudan a la detección de movimientos y cambio de posiciones por medio de la captación de rayos infrarrojos irradiados por el cuerpo humano en un área de 8*8 pixeles. Su principal ventaja es que puede ser usado tanto en luz como en oscuridad y detecta el cuerpo humano tanto en movimiento como en reposo, mientras que la desventaja es que no detecta objetos que sean indicio ASSB y no detecta las respiraciones [8]; también se realizó un estudio para la detección de respiraciones por medio de múltiples cámaras infrarrojas basándose en puntos de referencias faciales. Aunque esta tecnología alcanzó una tasa de efectividad del 71% en la detección de apneas, es necesario desarrollar estrategias que aumenten su precisión [7].

Uso de Cámaras e IA

Una de las innovaciones sobre la vigilancia de movimientos, y sobre la cual se basará esta investigación es el uso de cámaras e IA. Las cámaras permiten no solo la detección de movimientos sino también el análisis de la temperatura en imágenes 2D y 3D; esta tecnología ya se ha utilizado en la detección de parálisis cerebral en bebés mediante la identificación de movimientos anormales debido a que este sistema utiliza una cámara que apunta el bebé y una computadora que procesa la información capturando patrones de puntos en 2D en las extremidades, cabeza y torso del bebé. Estos patrones se utilizaron para entrenar la IA, diferenciando entre movimientos de riesgo y movimientos normales.

Una vez entrenada la IA es capaz de detectar y clasificar los movimientos en tiempo real [24], y que a diferencia de los proyectos con sensores, infrarrojo y ultrasonido, Adewopo et al. [10] señalan que la detección de patrones por medio de cámara se ven afectados por diversos factores como son el color de fondo, color de entorno, datos para entrenamiento de la inteligencia artificial y sobre todo, la capacidad de cómputo para realizar el proceso en tiempo real lo cual convierte esta tecnología en un reto para su producción en masa y a bajo costo.

El autor antes mencionado junto con su equipo, trabajaron en cuatro modelos de algoritmos para un sistema de seguridad física para el bebé en un hogar inteligente, llevando a cabo

reconocimiento de acciones por medio de cámaras inteligentes, logrando un *accuracy* entre el 80 y el 99 % de efectividad, crearon un modelo basado en convolución LSTM (Long Short-Term Memory) e I3D, y a diferencia de una red convolucional normal, la LSTM están diseñadas para manejar datos secuenciales que, como en este caso, se utilizó para detección de acciones por medio de secuencia de videos. Una gran desventaja de este método es la alta demanda de recursos computacionales.

Zhu et al. [14] realizaron un programa llamado *CribNet* y un conjunto de datos llamado *CribHD*, ambos se enfocan en la detección de objetos que se encuentran cerca del bebé y que pueden ser un riesgo para una muerte por asfixia, principalmente juguetes y mantas. Utilizaron diversos modelos como FSEG, SALL y el modelo *Mediapipe* (detección de rostros). El *CribHD* es una base creada en tres secciones: *CribHD-T* con 1,000 imágenes de juguetes, *CribHD-B* con 500 imágenes segmentadas de mantas y *CribHD-C* con 120 imágenes utilizando un muñeco, juguetes y mantas, y después de probar varios modelos como *YOLOv8*, *Detectron2* y *Yolact*, el *YOLOv8* fue el que tuvo mejores resultados obteniendo un mAP-50 promedio del 87% en las bases *CribHD* T y B.

En la evaluación de *CribHD-C* se evaluaron los modelos finales obtenidos obteniendo un 80% de *accuracy* en la identificación de juguetes peligrosos; también se señala que los estudios actuales se centran principalmente en el comportamiento del bebé o posturas, pero no abordan las identificaciones de situaciones peligrosas.

Bharati [11] diseñó un prototipo utilizando redes convolucionales para la detección de posiciones al dormir con un maniquí, las categorías que utilizó fueron boca arriba, transición a pose de riesgo (boca abajo) y posición de riesgo, obteniendo una base inicial de datos de 9,000 imágenes y realizando una estandarización de imágenes a 100*100 imágenes a blanco y negro, además de la realización de rotaciones de imágenes terminando con 18,000 de ellas. Después de varias pruebas, el mejor modelo pesaba 500kB y obtuvo el 100% de precisión con una respuesta de mensaje de tres segundos a una aplicación móvil, esta investigación generó ideas para el desarrollo de la metodología.

Kliker et al. [13] desarrollaron un prototipo para la detección de diversos factores de riesgo donde implementan el modelo *PoseNet*, que es un modelo de IA para la detección de posiciones por medio de puntos clave en el cuerpo y que, a diferencia la presente investigación desarrollada por nosotros que evalúa la imagen directamente, también detecta las respiraciones por medio de una cámara infrarroja, enfocándose en el cambio de temperatura debajo de la nariz por inhalar y exhalar, sin embargo, es un prototipo que aún está en fase experimental y no señalan ningún valor de precisión del sistema en la detección de posiciones o respiraciones.

Joseph C. Afonne [12] realizó una investigación para la detección de estrellas de mar corona de espinas utilizando *YOLO*, en su investigación probó diversos modelos de *YOLO* obteniendo resultados de *accuracy* superiores al 97% y un mAP de 99.3% tanto en 100 épocas como 200, señalando que los modelos *YOLO* es muy eficaz para la detección de objetos.

Una gran ventaja de la implementación de cámaras e IA es que no es un sistema invasivo a diferencia de los otros métodos para la detección de movimientos, por lo que es un método apto para el bebé, desafortunadamente no se cuentan con estadísticas de fiabilidad [9].

DELIMITACIÓN DEL OBJETO DE ESTUDIO

Planteamiento del Problema y Pregunta de Investigación

En la actualidad se han creado equipos capaces de detectar algunas variables que son factores de riesgo de muerte de cuna, se han creado equipos que implementan sensores físicos que pueden colocarse en la piel como en las prendas de vestir pero que, sin embargo, tienen contacto directo con la piel o utilizan radiaciones que pueden ser nocivas para la piel o salud del bebé. También se encontraron equipos que emplean luz infrarroja, uso de ondas de ultrasonido y cámaras que implementan IA, sin embargo, aún están en fase de prueba y/o llegan a ser muy costosos, además que en el mercado mexicano no contamos con estos equipos para uso doméstico y que sean accesibles. En las investigaciones revisadas, la mayoría de los equipos no detectan varios factores de riesgo de forma paralela, por lo que es necesario desarrollar un sistema que detecte



factores de riesgo para SMSL, además de factores de ASSB y otros factores como apneas prolongadas. Es por ello que se plantea la siguiente pregunta:

¿Cómo diseñar dos programas donde el primero detecte los factores de riesgo (SMSL y ASSB) con una alta eficiencia (*accuracy* mayor a 85% en un bebé y un maniquí), y otro que detecte puntos rojos para medir las respiraciones y apneas con un *accuracy* mayor a 85%?

Hipótesis Predictorias

Hi1.- Se puede crear un programa para detección de posiciones que emplea IA con imágenes de un maniquí con un *accuracy* superior al 85%.

Ho1.- No se puede crear un programa para detección de posiciones que emplea IA con imágenes de un maniquí con un *accuracy* superior al 85%.

Hi2.- Existen diferencias significativas en el *accuracy* en los valores obtenidos en el programa de detección de posiciones entre la base de datos con maniquí y un bebé.

Ho2.- No existen diferencias significativas en el *accuracy* en los valores obtenidos en el programa de detección de posiciones entre la base de datos con maniquí y un bebé.

Hi3.- Es posible crear un programa que detecte puntos rojos en el abdomen de un maniquí y un bebé con un *accuracy* mayor a 85%.

Ho3.- Es posible crear un programa que detecte puntos rojos en el abdomen de un maniquí y un bebé con un *accuracy* mayor a 85%.

Objetivos

Objetivo General

Desarrollar 2 programas de IA donde el primero detecte factores de riesgo asociados con la muerte de cuna como posiciones físicas de riesgo (lateral, boca abajo y cuando tenga la cabeza tapada) y presencia de objetos que puedan provocar una muerte por asfixia, mientras que el segundo detecte 2 puntos rojos en el abdomen que ayude a detectar las respiraciones y apneas. Se buscará que tengan un *accuracy* mayor a 85% tanto teórico (imágenes de maniquí) como en lo experimental (imágenes de bebé) para poder ser implementados en un futuro en un sistema integral.



Objetivos Específicos

- Desarrollar y probar un programa basado en IA que reconozca e identifique posiciones de riesgo de muerte de cuna y por asfixia con un *accuracy* mayor a 85% tanto en fase de entrenamiento como de validación.
- Comparar los resultados entre las fases de entrenamiento y validación con una fase de prueba (tercera fase) con una base de datos de un bebé real aplicando el modelo final obtenido en las primeras 2 fases.
- Diseñar un programa basado en IA para la detección de puntos rojos con un *accuracy* mayor a 85% en las tres fases.

JUSTIFICACIÓN

A diferencia de los sistemas tradicionales basados en sensores físicos, que pueden presentar riesgos de contacto directo o limitarse a un hardware específico, la tecnología basada en IA es más flexible y segura. La tecnología basada en IA es segura físicamente para el bebé además de tener un gran potencial de integrar mejoras a través de actualizaciones de software, aumentando su flexibilidad, confiabilidad y una continua mejora con el mismo hardware.

En este trabajo se busca desarrollar 2 programas que implementen IA para la detección oportuna de factores de riesgo y la falta de respiraciones, se tomarán fotografías por medio de una cámara por lo que es un método seguro para la salud del bebé; dichos programas serán diseñados para poder ser implementados en proyectos futuros y llevarlo a un sistema integral.

Es importante mencionar que todo el diseño del sistema es original, desde la parte de software (programación de IA y otros programas) como hardware (set-up), se espera que ambos programas logren un buen *accuracy* tanto con el maniquí como de bebé real con el fin de poder desarrollar un programa que pueda ser implementado en un proyecto a futuro para patente.



MATERIALES Y MÉTODOS

Diseño del Estudio

Este proyecto cuenta con un enfoque cuantitativo con diseño experimental. Es cuantitativo debido a que se medirá el porcentaje de efectividad de los programas y las variables dependientes son magnitudes; así mismo tiene un diseño experimental, porque tiene varias etapas donde hay desarrollo original de base de datos, diseño de set up y programas para el reconocimiento de las posiciones de riesgo y detección de respiraciones [25]; igualmente esta investigación también es de carácter aplicado porque se obtendrán 2 programas que serán implementados para la creación de un sistema integral a futuro.

Unidad de Estudio

Tamaño de Muestra

Shahinfar [26] realizó una investigación para encontrar el número ideal de imágenes para un entrenamiento de red neuronal en el cual se utilizaron 20, 50, 150, 500 y 1000 imágenes por clasificación encontrando que la exactitud, precisión y el número de falsos positivos y negativos mejoran logarítmicamente incrementando el número de imágenes en un conjunto de entrenamiento, sin embargo, cuando el número de imágenes sobrepasa las 500 imágenes las mejoras fueron insignificantes, dando un intervalo de 150 a 500 de imágenes para lograr una precisión razonable.

Tal como menciona el autor [26], el número de imágenes a utilizar en una IA dependerá del tipo de investigación y de las estadísticas arrojadas durante el entrenamiento para encontrar la mejor tasa de confiabilidad. Ballesteros et al. [27] utilizaron 25,000 imágenes y señalan que la eficiencia de un sistema de IA depende mucho de la calidad y estandarización de las imágenes en cada una de las clasificaciones que se utilicen.

Con base a lo anterior se decidió realizar una base de datos inicial con 400 imágenes en cada categoría. De las fotografías tomadas para la fase de prueba del bebé solo se seleccionarán 100 de cada categoría, debido a que solo evaluarán el modelo final del programa.

Materiales

Para el diseño del set up, se utilizaron los siguientes materiales:

- Tubería PVC hidráulico de media pulgada
- Muñeco maniquí con características parecidas a un recién nacido
- Sabana color azul claro
- Tabla de madera de 120 cm por 90 cm, simulando las dimensiones promedio de las camas de cunas actuales
- Cámara digital para computadora alámbrica
- Computadora propia para realizar la programación
- Ropa para bebé de 0 meses
- Bolsa Foley
- Juguetes para bebés
- Sabana gris

Definición de Variables

Tanto las variables dependientes como independientes son presentadas en la Tabla 1 junto con sus indicadores y escalas de medición.

Tabla 1: Definición de variables dependientes e independientes.

Variable	Definición conceptual	Definición operacional	Indicadores	Escala de medición
Independiente: Fotografías del maniquí para programa de detección de posiciones	Fotografías tomadas en diferentes posiciones y tonalidades de iluminación.	Se tomarán fotografías en diversas posiciones e iluminaciones para crear una base de datos robusta y que contenga la información necesaria para poder entregar un buen modelo de programa	Se tomaron fotografías con una resolución estandarizada basándonos en las categorías que se trabajarán	400 fotografías de 640*480 pixeles por categoría
Independiente: Fotografías del bebé para programa de detección de posiciones	Fotografías tomadas en diferentes posiciones y tonalidades de iluminación	Se tomarán fotografías en diversas posiciones e iluminaciones para crear una base de datos robusta y que contenga la información necesaria para realizar una fase de prueba	Se tomaron fotografías con una resolución estandarizada basándonos en las categorías que se trabajarán	100 fotografías de 640*480 pixeles por categoría

Independiente: Fotografías para programa de detección de puntos rojos	Fotografías tomadas en diferentes posiciones estando boca arriba para ubicación de puntos rojos	Se creará una base de datos con imágenes en posición boca arriba con diversas iluminaciones y ubicaciones tanto del maniquí como del bebé	Una vez procesadas nos arrojará el programa un resultado, determinando si las fotografías fueron las adecuadas	500 imágenes de 640*480 pixeles
Dependiente: <i>Accuracy</i> y <i>Loss</i> en programa de detección de posiciones.	Es la exactitud y pérdida del programa en la fase de entrenamiento, validación y prueba	Durante la fase de entrenamiento, validación y prueba, el sistema medirá la efectividad del programa proporcionando el <i>accuracy</i> y el <i>loss</i> de las tres fases.	En cada época dará resultados del procesamiento y serán plasmadas en una tabla para su interpretación. Además de una tabla de correlación en la fase de prueba	Porcentaje de 0 a 100
Dependiente: <i>Accuracy</i> , <i>Recall</i> , <i>mAP50</i> y <i>Loss</i> en programa de detección de puntos rojos	Es la eficacia y pérdida del programa en la fase de entrenamiento, validación y prueba	Durante la fase de entrenamiento, validación y prueba, el sistema medirá la efectividad del programa proporcionando métricas de la precisión y pérdida de las tres fases	En cada época dará resultados del procesamiento y serán plasmadas en una tabla para su interpretación. Además de una tabla de correlación en la fase de prueba	Porcentaje de 0 a 100

Nota: Definición de las variables dependientes como independientes utilizadas en esta investigación, definición operacional, como es que se evaluarán y su escala de medición. En ambos programas las variables independientes serán las imágenes utilizadas como base de datos en cada uno, mientras que las dependientes son *Accuracy* y *Loss* obtenidos en los programas (diseño propio).

Técnicas y Procedimientos

El modelo que se implementó en el programa de detección de posiciones es una red neuronal de tipo perceptrón multicapa convolucional para pronosticar posiciones de riesgo para el neonato y la eficiencia del modelo se midió por medio de entropía cruzada generando así una curva de precisión durante el proceso de entrenamiento.

En una segunda fase se generó una curva similar durante el proceso de validación del modelo, para el proceso de entrenamiento se utilizaron 80% de las muestras y el 20% para

validación, estos porcentajes son un estándar de referencia [14]. Durante el proceso de prueba se generaron curvas ROC y se construyeron las matrices de falsos positivos y falsos negativos.

El proyecto se divide en 7 etapas, las cuales son: 1) Diseño de set up, 2) Creación de base de datos, 3) Diseño de instrumento, 4) Validez de instrumento, 5) Creación de programa de IA para la detección de factores de riesgo, 6) Aplicación de programación con fotos de bebé, y 7) creación de programa para la detección de respiraciones con maniquí. A continuación, serán explicadas.

Etapas 1. Diseño de Set-up:

Primeramente se tomaron en cuenta las medidas estándar de una cuna normal, después se realizó un diseño propio que consiste en una tabla de madera de 60 x 100 cm, una manta y tubería PVC hidráulico que sostiene una cámara digital lo más céntrico posible a la cama; en la Figura 3 se muestra la organización de los materiales utilizados, y también se muestra el funcionamiento del sistema con un armazón diseñado para colocar la cámara en un punto específico para entregar una imagen centralizada.

Figura 3: Primer set up de cuna con cámara y la obtención de imagen por computadora.



Nota: Primer diseño del set-up utilizando una laptop para observar el funcionamiento, el maniquí con la prenda diseñada y tubería PVC para ubicar la cámara. En la prenda del bebé se colocaron 2 puntos rojos a la altura del estómago para la posterior detección de respiraciones (fuente propia).

Se colocaron en la prenda de vestir dos círculos rojos en la parte abdominal y un cuadro verde debajo de la nuca del maniquí, esto con el fin de ayudar a los programas a la detección de las posiciones.

Etapas 2: Creación de Base de Datos

Se tomaron fotografías en las cinco diferentes clasificaciones las cuales se detallan en la Tabla 2. En un inicio se tomaron esas fotografías a una resolución de 1280*780 pixeles, sin embargo, la falta de recursos computacionales no permitió la conclusión de la programación por lo que se cambió a una resolución de 640*480 pixeles.

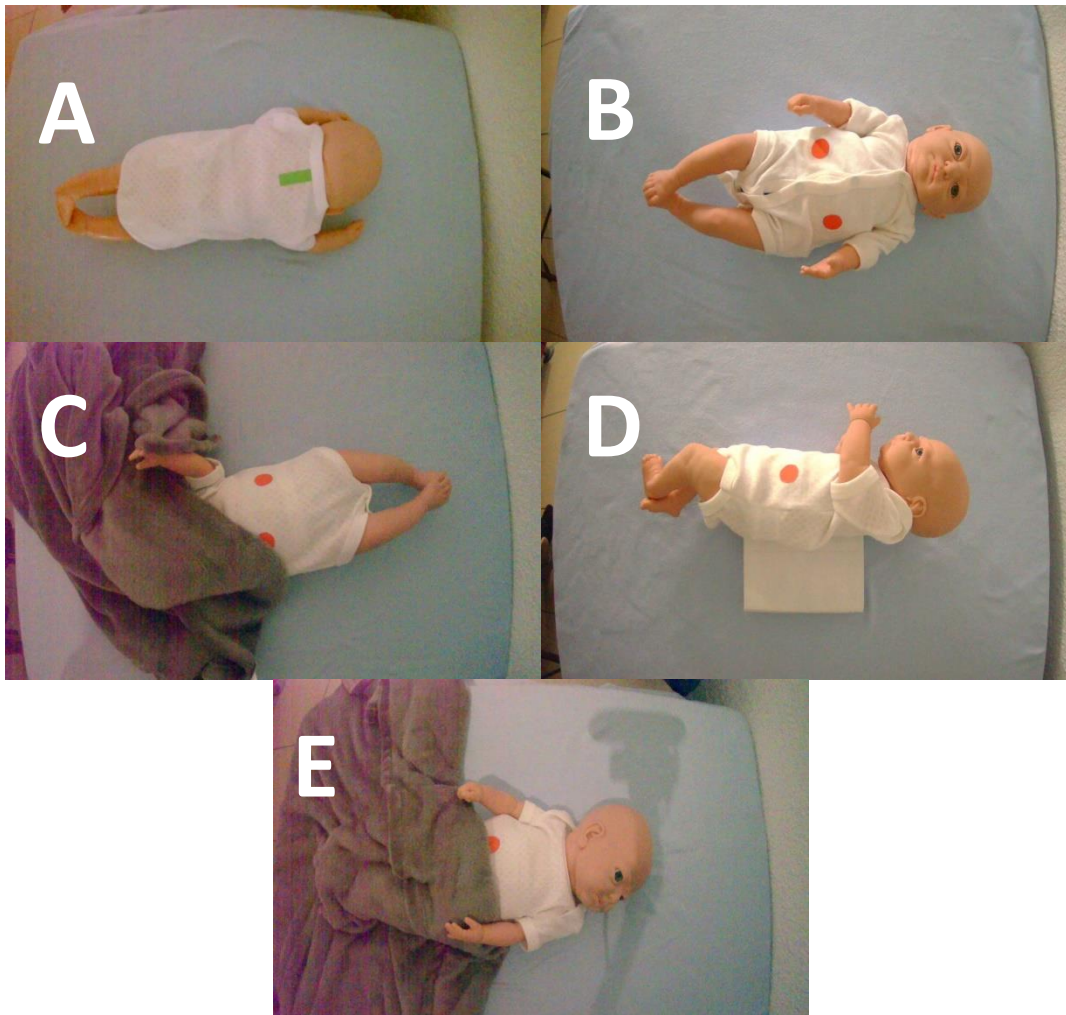
Tabla 2: Categorías a detectar con IA.

Categoría	Descripción	No. de muestras
Normal	Posición normal del maniquí (decúbito supino), simulando una posición normal de reposo	400
Lateral	Posición lateral del maniquí	400
Tapado	Tanto en posición lateral como normal existe la presencia de una cobija o manta tapando el abdomen del bebé sin pasar el nivel de las costillas	400
Boca abajo	Posición de mayor riesgo (decúbito prono), se puede detectar tanto con cobija como sin cobija	400
Cabeza tapada y objetos	Cualquier posición del bebé en la que una cobija u objeto tape la cabeza del bebé o sea un riesgo	400

Nota: Descripción de las posiciones que se trabajaron en el programa de detección de posiciones y el número de muestra de cada una de ellas (diseño propio).

La Figura 4 da un ejemplo de imagen de cada una de las categorías. También se procuró obtener en cada imagen diferentes iluminaciones, se modificaba la posición de las extremidades, cabeza, objetos y ubicación de puntos rojos. El punto verde ubicado en la parte trasera de la prenda en la imagen A) ayuda a la IA a detectar más fácilmente la categoría “boca abajo”.

Figura 4: Fotografías de las 5 categorías.



Nota: Las imágenes muestran un ejemplo de las categorías: A) Boca abajo, B) Normal, C) ASSB, D) Lateral y E) Normal tapado. También se puede apreciar diferentes iluminaciones en las imágenes (fuente propia).

Se utilizó el lenguaje de programación *Python* en la plataforma *Google Colab* debido a que es una plataforma gratuita de *Google* que ofrece recursos computacionales en línea para facilitar el desarrollo de programas y su ejecución. Para ayudar en la fase de análisis, se diseñó el primer programa para realizar una redimensión de imágenes para facilitar el proceso, dichos cambios fueron realizados varias veces hasta encontrar un tamaño óptimo.



Posteriormente se realizó el segundo programa para seleccionar 100 imágenes aleatorias de cada una de las categorías de la base de datos y realizar una copia de dichas imágenes para rotarlas 90°, sumándolas a la base de datos. Este proceso se repitió para obtener 100 imágenes aleatorias a 180° y 100 a 270°, obteniendo los programas tres y cuatro respectivamente. Estos programas ayudaron al aumento de base de datos, mejorando la efectividad del programa para la detección de factores de riesgo de muerte de cuna.

Etapas 3: Diseño de Instrumento

Debido a que este proyecto consiste en el procesamiento de imágenes, se realizó una lista de cotejo con ocho ítems (pasos a seguir) para la estandarización y mejor procesamiento de las imágenes. Cabe mencionar que no se basó en ningún parámetro, y esta metodología fue propia de esta investigación. Los ítems inicialmente fueron:

1. Redimensionar fotos para estandarizar tamaño
2. Detección de fondo de imagen
3. Detección de bordes
4. Detección de esqueleto
5. Análisis de Histogramas
6. Normalización de imágenes
7. Rotación de imágenes
8. Reflexión de imágenes

Este instrumento fue aplicado mediante un cuestionario digital enviado a través de medios electrónicos, primero fue evaluado por expertos tanto en validez como en contenido.

Etapas 4: Validez del Instrumento

Con el fin de validar el contenido del instrumento, se revisaron cada uno de los ítems mediante el análisis de validez de contenido por medio de juicio de expertos. Dicho instrumento fue diseñado en *Google Forms*, la cual es una plataforma gratuita de *Google* que permite diseñar encuestas para poder ser compartidas de manera digital. Teniendo el cuestionario terminado, se envió a varios expertos en el área de IA.

Primeramente, se llevó a cabo una fase de evaluación de expertos. Se creó un cuestionario para evaluar la pertinencia, la relevancia y redacción de los ítems que conforman el instrumento. La escala de medición fue de: “No”, “Medianamente” y “Si” detallada en la Tabla 3, no obstante, para fines de validación de contenido (CVI) se colocaron valores de cero a uno respectivamente. Además de esta evaluación, se colocó un apartado en cada ítem para que los expertos dieran sus opiniones.

Tabla 3: Encuesta realizada en cada uno de los ítems de la fase de evaluación de expertos.

Pregunta	Respuestas	Valor
¿Considera que el ítem es pertinente y relevante?	Si	1
	Medianamente	0.5
	No	0
¿Considera que el ítem es perteneciente y relevante?	Si	1
	Medianamente	0.5
	No	0
Observaciones	Respuesta libre	

Nota: En cada ítem se realizaban las preguntas de si era pertinente y perteneciente al motivo de la investigación, y en base a la respuesta se determinaba un valor para poder ser procesada (diseño propio).

Como estrategia para evaluar y clasificar las evaluaciones, se consideró que los ítems con valores de 0 - 0.5 son irrelevantes y serían eliminados de la lista, de 0.5 - 0.8 serán considerados aceptables dependiendo los comentarios y, los que tuvieran calificación de uno serían aceptables (relevantes).

Una vez obteniendo los resultados de la fase de evaluación de expertos y realizar los cambios sugeridos a la lista de cotejo, se llevó a cabo otro cuestionario para una fase de Juicio de expertos en el que se midió el nivel de pertinencia y redacción de cada ítem.

Para ello se llevó a cabo una escala de Likert con cuatro valores (1=Inaceptable, 2=Deficiente, 3=Aceptable y 4=Eficiente) señalando el nivel de pertinencia y comprensión de cada ítem, mismo que sirvió para la evaluación.

Además de analizar en nivel de pertinencia y de comprensión, se colocó un apartado de sugerencias como en la fase anterior. Esta fase de juicio de expertos se llevó a cabo con la finalidad de determinar cuáles eran las acciones a realizar a la base de datos para llevar a cabo esta investigación. La descripción de las preguntas de cada uno de los ítems se detalla en la Tabla 4.

Tabla 4: Preguntas para cada ítem de la fase Juicio de expertos.

Análisis de ítems	Respuestas
Determine el grado de pertinencia del ítem	1) No es pertinente. El ítem no contribuye a evaluar el propósito y dimensión del instrumento en ningún aspecto. Puede ser eliminado. 2) Bajo nivel de pertinencia. El ítem hace una contribución superficial a la evaluación del propósito y dimensión del instrumento. 3) Aceptable grado de pertinencia. El ítem contribuye a evaluar el propósito, dimensión y/o constructo teórico del instrumento. 4) Alto nivel de pertinencia. El ítem contribuye a evaluar en un alto grado el propósito, dimensión y/o constructo del instrumento.
Determine el grado de redacción del ítem	1) No es comprensible. El ítem no es comprensible por los potenciales usuarios del instrumento. 2) Bajo nivel de comprensión. El ítem debe ser mejorado en al menos la mitad de sus componentes para un claro entendimiento. 3) Aceptable nivel de comprensión. El ítem requiere algunas mejoras superficiales para que quede bien. 4) Alto nivel de comprensión. El ítem es altamente comprensible por los potenciales usuarios y sigue las normas gramaticales de la lengua.
¿Tiene alguna sugerencia para mejorar el ítem?	Respuesta libre

Nota: En la encuesta virtual se presentaba en cada ítem estas 3 preguntas de análisis y se pedía que se contestara con un valor de 1 a 4, a excepción de la última pregunta que era abierta (diseño propio).

Una vez obtenidas las respuestas de juicio por expertos se procedió a revisar la adecuación de los ítems mediante Índice de Validez de Contenido (IVC) [28], el cual proporciona una

magnitud que informa entre la proporción de jueces que manifiestan una calificación positiva (3 y 4) y una calificación negativa sobre un ítem (0 y 1). Esta puntuación ayudará en la toma de decisiones en cuanto a modificar el ítem o eliminarlo. La fórmula para el cálculo del IVC para cada ítem fue realizada con la siguiente fórmula:

$$IVC = \frac{\sum n}{N}$$

Siendo n las calificaciones obtenidas por los expertos (0 a 1 para la primera fase y 1 a 4 para la segunda) y N es el número de expertos. Para el cálculo de S-CVI/Ave se utilizó la siguiente fórmula:

$$S - CVI/Ave = \frac{\sum(I - CVI)}{N}$$

El cuál es el resultado de la suma del CVI de cada ítem entre el número de expertos. También un parámetro a analizar fue S-CVI/UA con la siguiente fórmula:

$$S - CVI/UA = \frac{N^{\circ} \text{ items con } I - CVI \geq 0.67}{N}$$

La diferencia es que este solo considera ítems que tuvieron un CVI mayor a 0.67. Se considero que 0.67 es un valor aceptable en la que mínimo 2 jueces de 3 validaban el ítem.

Etapas 5: Creación de Programa de IA para la Detección de Factores de Riesgo

Posteriormente a los programas para aumentar el número de imágenes se diseñó un quinto programa para la selección de imágenes para ser usadas tanto en la fase de entrenamiento y la fase de prueba. Como se mencionó en apartados anteriores, se optó por utilizar 80% de las imágenes en la fase de entrenamiento y el restante 20% en la fase de prueba.

Para llevar a cabo la selección se llevó a cabo un enlistado el total de imágenes (3,500), después el mismo programa seleccionaba aleatoriamente el 80% de la lista de imágenes de cada categoría (560 imágenes por carpeta o categoría) y creaba un archivo de texto con ese enlistado. Con el restante 20% se creó la lista de validación.



Después se realizó un programa de IA convolucional aplicando ResNet. Yani et al. [29] realizaron diversas pruebas con diversos valores de batch y epoch, además de diferentes porcentajes de muestra en entrenamiento y validación. El menor valor de *loss* se obtuvo con un batch mínimo de ocho y 16, y con el mayor número de épocas posibles. Además la selección del 90% de datos para entrenamiento y 10% para validación fueron los que obtuvieron mejores resultados. Sin embargo el uso del 80% de las muestras en fase de entrenamiento y 20% en fase de validación o prueba es un estándar utilizado en varias investigaciones [14].

Al inicio del programa se analizaba cada imagen de los archivos de texto de entrenamiento y validación, se llevaron a cabo varios intentos de análisis y modificaciones buscando el mejor resultado obteniendo el mejor utilizando 25 épocas de análisis, un batch de 16 imágenes y una tasa de aprendizaje de $1e-5$. Debido a que eran 2,800 imágenes de entrenamiento y 700 imágenes para validación con un batch de 16 en cada uno se realizaron 175 iteraciones para el entrenamiento y 43 para la validación.

Durante cada época el modelo ajusta sus pesos (parámetros) en función a los datos que procesa. Al inicio de cada época el modelo genera predicciones sobre los datos de entrenamiento y calcula el *accuracy* y *loss*, utilizando *loss* para ajustar sus parámetros y crear otro modelo mejorado en la siguiente época.

Durante todo el proceso de análisis se llevaron a cabo cambios en la base de datos como en la programación para obtener un mejor resultado. Una vez terminado el mismo programa guarda el modelo que mejores resultados obtuvo. El modelo final se realizó en 8 horas, 9 minutos y 12 segundos y fue implementado en la fase de prueba tanto para fotos de un maniquí como fotos del bebé.

Etapas 6: Implementación de Programación con Fotos de Bebé (Fase de Prueba)

Una vez terminado el programa de entrenamiento y validación se implementó el modelo con mejores resultados en la fase de prueba, observando la efectividad del modelo con imágenes de un bebé real en situaciones reales.

Se llevó a cabo un consentimiento informado hacia los padres de un bebé de cuatro meses, el cual se puede observar en la sección de anexos. Para la fase de prueba se utilizó el set-up para tomar fotos al bebé en diversas posiciones de las cinco categorías que se trabajan en este proyecto y un video de 30 segundos respirando en posición boca arriba. Cabe mencionar que siempre se llevaron acciones con la presencia y apoyo de los padres en todo momento.

La Figura 5 es una fotografía autorizada del bebé en posición Normal para ser utilizada en este documento con el fin de corroborar la actividad descrita. Cabe mencionar que se tomaron extremas precauciones para la toma de fotografías de las categorías Boca abajo y ASSB sin afectar al bebé.

Figura 5: Ejemplo de imagen del bebé en posición Normal.



Nota: Esta imagen fue autorizada por los padres para ser usada en este documento. Varias de las posiciones realizadas por el bebé fueron hechas en el maniquí para un mayor realismo y asertividad en la fase de entrenamiento (fuente propia).



Etapas 7: Creación de Programa para la Detección de Respiraciones

Primeramente, se llevó a cabo la toma de 500 imágenes del maniquí con diferentes iluminaciones y posiciones para la identificación de los puntos en el abdomen, después se extrajeron 50 imágenes del video del bebé respirando para usarlas en la fase de prueba.

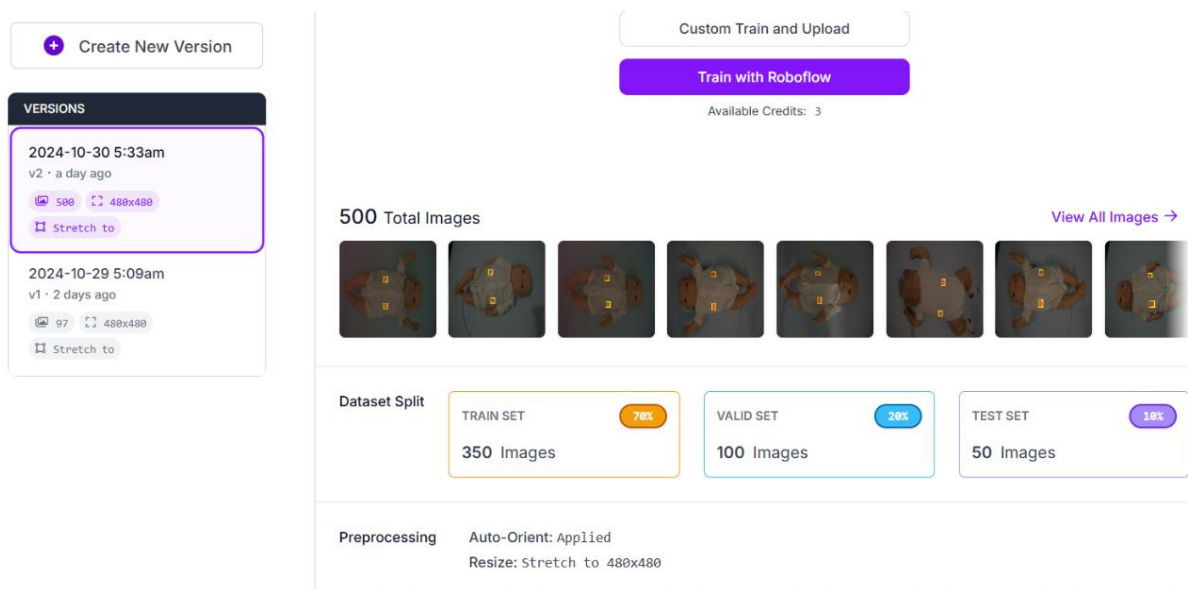
Para facilitar la detección de los puntos rojos en el abdomen en tiempo real se optó por desarrollar un programa basado en *YOLO*, ya que es popular por la rapidez y precisión con la que detecta objetos en tiempo real.

Para poder utilizar *YOLO* se requirió etiquetar las imágenes (proceso también conocido como *annotation* o *labeling*) que es asignar etiquetas en áreas específicas de las imágenes, en este caso los puntos rojos. Para ello se utilizó la página de internet *Roboflow* que permite realizar esta actividad. Cabe mencionar que es una página “freemium” lo que significa que se puede usar gratuitamente y solo se utilizaron imágenes del maniquí.

El primer paso es ingresar las imágenes a trabajar para después redimensionarlas, en esta fase se seleccionó 480*480 píxeles. Posterior a eso se etiquetaron cada una de las 500 imágenes seleccionando los 2 puntos abdominales con la etiqueta “Puntos_rojos”, después del proceso de etiquetado se debe seleccionar el porcentaje de imágenes que serán utilizadas para la fase de entrenamiento, validación y prueba, en el que se seleccionaron los porcentajes 70, 20 y 10 respectivamente.

Terminando el proceso se puede descargar un archivo comprimido que puede ser utilizado en los programas que usan *YOLO*, en el archivo comprimido vienen las imágenes categorizadas por la fase donde se utilizarán y un archivo de la ubicación de las etiquetas de cada una de las imágenes. La Figura 6 es una captura de pantalla finalizando el etiquetado de las 500 imágenes y la selección de 350 para la fase de entrenamiento, 100 para validación y 50 imágenes para prueba.

Figura 6: Etiquetado de 500 imágenes en Roboflow.

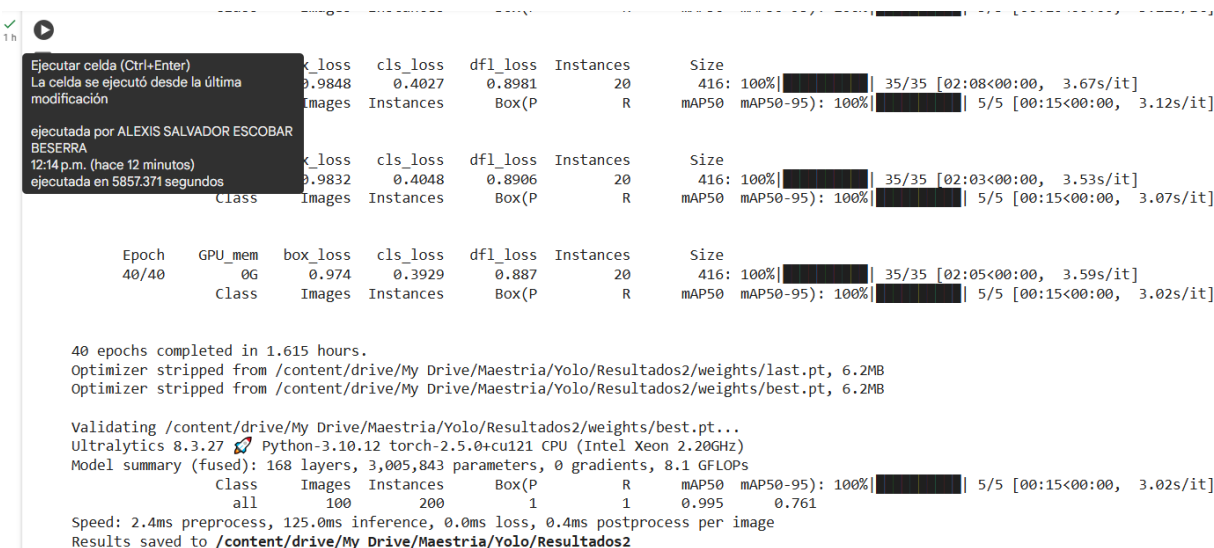


Nota: Se puede apreciar las 2 etiquetas en las imágenes del bebé como puntos amarillos sobre los puntos rojos del abdomen, al finalizar el etiquetado se seleccionaba las imágenes para las fases de entrenamiento, validación y prueba, seleccionando los porcentajes 70%, 20% y 10% de las imágenes respectivamente. Por último, se hizo un redimensionamiento de imágenes a 480*480 píxeles para un procesamiento más rápido. Estas acciones se llevan a cabo dentro de la plataforma (fuente propia).

Después se llevó a cabo una programación en *Colab* para crear el programa de detección de puntos rojos, se llevaron a cabo varias pruebas dando el mejor resultado utilizando 40 épocas con un tamaño de batch de 10 imágenes, por lo que se llevaron a cabo 35 iteraciones (o pasos) en la fase de entrenamiento y cinco para la fase de validación en cada época.

El proceso de las 40 épocas se llevó a cabo en una hora con 37 minutos y 37.37 segundos entregando métricas de desempeño cada época y realizando una fase de validación con resultados favorecedores. La Figura 7 es una captura de pantalla del programa en *Colab* donde se muestra la conclusión de la época (epoch) 40 con 100% de *accuracy* y sus 35 iteraciones además de la fase de validación. Este proceso también entrega un modelo final que puede ser implementado para la fase de prueba o aplicarlo en algún otro proyecto.

Figura 7: Proceso de análisis de Yolo.



Nota: Culminación del proceso de entrenamiento y validación del programa de detección de puntos rojos en *Colab*. Los resultados obtenidos fueron satisfactorios desde antes de la época 40 y finalizo en una hora con 37 minutos. Foto propia.

Consideraciones Éticas

Esta investigación se someterá a una revisión del comité de ética de la Facultad de Medicina y Ciencias Biomédicas de la Universidad Autónoma de Chihuahua. Los procedimientos de esta investigación cumplen con la Declaración de Helsinki en materia de la investigación en humanos y su dignidad. Para realizar la toma de fotografías en un bebé se llevó a cabo un consentimiento informado considerando la Ley general de Protección de Datos Personales en Posesión de Sujetos Obligados (LGPDPPO), la Ley general de Protección de Datos Personales en Posesión de los Particulares, el Reglamento de Ley General de Salud en Materia de Investigación en Salud, el código civil federal, Ley Federal de Protección de Datos Personales, el Código Penal Federal y la Ley General de los Derechos de Niños, Niñas y Adolescentes.

Se asegurará la confidencialidad de la información personal tanto del bebé como de los padres mediante el anonimato de sus datos, el almacenamiento seguro de la información con acceso restringido únicamente al investigador, esto mediante autorización de los cuidadores asegurando la participación voluntaria y podrá ser revocada en cualquier momento.

RESULTADOS

Introducción

En este capítulo se presentan los datos obtenidos del análisis del instrumento, así como el análisis de la información obtenida por los 2 programas (Detección de respiraciones y detección de puntos rojos). Para el instrumento se muestra la prueba de normalidad y confiabilidad, después se detallarán las bases de datos utilizadas y por último se dará un análisis en tres niveles: univariado, bivariado y multivariado de los programas mencionados.

Análisis de Instrumento

De la encuesta enviada a diversos expertos se recibió respuesta de tres expertos, los cuales debían tener las características de tener grado de doctorado y contar con más de tres años de experiencia en el área de IA.

Validación de Contenido

En la revisión inicial de la lista de cotejo, el juez número 2 realizó más recomendaciones y críticas respecto a la lista de cotejo. Todos los ítems fueron modificados para una mejor comprensión basándonos en los comentarios de los expertos. La

Tabla 5 señala los resultados obtenidos y las modificaciones que se realizaron a los ítems, cabe mencionar que esos ítems finales fueron los que se utilizaron para la fase de Juicio de expertos.

Tabla 5: Resultados de fase evaluación de expertos.

Ítem	Juez	Pertinencia	Redacción	Comentarios	Ítem final
1	1	Si	Si	No	Estandarizar tamaño de fotografías a 480 x 640 pixeles para que sea más rápido su análisis y detectar más fácilmente el maniquí.
	2	Si	Medianamente	Si	
	3	Si	Si	No	
2	1	Si	Si	Si	Detectar el fondo para convertirlo a un valor 0 (fondo
	2	Medianamente	Si	Si	

	3	Si	Si	No	negro) para diferenciarlo del maniquí, lo cual ayuda en la detección del esqueleto.
3	1	Si	Si	Si	Detección de bordes del maniquí para separarlo del fondo.
	2	Medianamente	Medianamente	Si	
	3	Si	Si	No	
4	1	Si	Si	Si	Una vez detectado el maniquí y diferenciado del fondo (convertido a valor 0) se debe convertir el maniquí a valor 1 (blanco) para aplicar la función "skeleton" y así detectar el esqueleto del maniquí.
	2	No	Si	Si	
	3	Si	Si	No	
5	1	Si	Si	No	Obtener un histograma de colores para eliminar colores indeseados o su sombra que provoquen ruido en la detección del maniquí.
	2	No	Si	Si	
	3	Si	Si	No	
6	1	Si	Si	Si	Normalizar un tamaño de cuadrado para la detección del maniquí.
	2	Medianamente	Medianamente	Si	
	3	Si	Si	No	
7	1	Si	Si	No	Rotación de imágenes para la detección del maniquí en diferentes posiciones para el entrenamiento de la inteligencia artificial.
	2	Si	Si	No	
	3	Si	Si	No	
8	1	Si	Si	No	Rotación de imágenes para la detección del maniquí en diferentes posiciones para el entrenamiento de la inteligencia artificial.
	2	Si	Si	No	
	3	Si	Si	No	

Nota: Resultados obtenidos en la fase de validación de contenido, se agregó descripción a cada uno de los ítems y se evaluaron los comentarios de los ítems que tuvieron opiniones (diseño propio).

Posterior a la obtención de los resultados se procede a realizar la fase de evaluación de contenido. Debido a que evaluaron tres jueces se consideró que un valor de 0.67 o más era

aceptable. Los ítems 4 y 5 fueron los que recibieron menos valoración, sin embargo, analizando los comentarios fue evidente que hacía falta una mejor comprensión de los ítems por lo que se decidió no eliminar a ninguno y modificarlos para un mejor entendimiento.

En la

Tabla 6 se detallan los resultados de pertinencia de la evaluación de contenido, en ella se muestra el CVI e I-CVI de cada ítem, dando un S-CVI/Ave (promedio de índices de validez) de 0.793 y un S-CVI/UA (índice de validez) de 1.

Tabla 6: Resultados de Pertinencia del instrumento

Ítem	J1	J2	J3	CVI	I-CVI
1	1	1	1	1	1
2	1	0.5	1	0.833	0.67
3	1	0.5	1	0.833	0.67
4	1	0	1	0.67	0.67
5	1	0	1	0.67	0.67
6	1	0.5	1	0.833	0.67
7	1	1	1	1	1
8	1	1	1	1	1

Nota: Resultados del nivel de pertinencia de cada uno de los ítems. Los más criticados fueron el 2, 3 y 6, por lo que fueron objeto de análisis para su eliminación (diseño propio).

La Tabla 7 contiene los resultados de redacción de la fase de validación de contenido y también se muestra el CVI e I-CVI de cada ítem, dando un S-CVI/Ave de 0.8762 y un S-CVI/UA de 1. En base a los comentarios nos dimos cuenta que no estaban bien explicados los ítems por lo que se modificaron todos agregando detalles de cada uno. Al igual que el resultado de pertinencia, los ítems 3 y 6 tuvieron comentarios, por lo que se analizó su posterior eliminación.

Tabla 7: Resultados de Redacción del instrumento

Ítem	J1	J2	J3	CVI	I-CVI
1	1	0.5	1	0.833	0.67
2	1	1	1	1	1
3	1	0.5	1	0.833	0.67
4	1	1	1	1	1
5	1	1	1	1	1

6	1	0.5	1	0.833	0.67
7	1	1	1	1	1
8	1	1	1	1	1

Nota: Los ítems 1, 3 y 6 recibieron comentarios siendo el 3 y 6 más criticados como en la fase de evaluación de pertinencia, se optó por su posterior eliminación (diseño propio).

Validación de Confiabilidad

Después de analizar resultados obtenidos por juicio de expertos y validación por IVC se consideró tener un valor mayor a 0.8 para que fuera pertinente un ítem, tanto el ítem 3, 5 y 6 se eliminaron debido a que no se consideraron relevantes para fines de esta investigación. Los ítems 2 y 4 recibieron modificaciones en su escritura para un mejor entendimiento y el ítem 1 fue relevante en todos los sentidos.

Por último, los ítems 7 y 8 se fusionaron debido a que consisten en el aumento de datos creando un nuevo ítem colocado como ítem 9. En la Tabla 8 se agregaron los resultados obtenidos y se realizaron las modificaciones finales a los ítems.

Tabla 8: Resultados de la fase Juicio de expertos.

Ítem	IVC de pertinencia	IVC de redacción	Modificaciones al ítem
1	1	1	Sin modificaciones
2	1	1	“Detección de fondo para diferenciarlo del maniquí, convirtiéndolo en fondo negro para ayudar en la detección del esqueleto.”
3	0.66	1	Eliminación de ítem
4	1	1	“Una vez detectado el maniquí y diferenciado del fondo (convertido a valor 0) se debe convertir el maniquí a valor 1 (blanco) para detectar el esqueleto.”
5	0.66	1	Eliminación de ítem
6	0.66	0.66	Eliminación de ítem

7	1	1	Fusión con ítem 8
8	0.66	0.66	Fusión con ítem 7
9			Nuevo ítem de la fusión del 7 y 8: “Aumento de datos aplicando técnicas de transformación rígida (traslación, reflexión y rotación)”

Nota: Resultados finales, los ítems 3, 5 y 6 fueron los que tuvieron menor calificación en comparación a los demás durante la fase de juicio de expertos, por lo que fueron eliminados (diseño propio).

La lista final quedo de la siguiente manera:

- Estandarizar tamaño de fotografías a 480 x 640 pixeles para que sea más rápido su análisis y detectar más fácilmente el maniquí.
- Detección de fondo para diferenciarlo del maniquí, convirtiéndolo en fondo negro para ayudar en la detección del esqueleto.
- Una vez detectado el maniquí y diferenciado del fondo (convertido a valor 0) se debe convertir el maniquí a valor 1 (blanco) para detectar el esqueleto.
- Aumento de datos aplicando técnicas de transformación rígida (traslación, reflexión y rotación).

Caracterización de Muestra

Como se mencionó en apartados anteriores las características de la muestra fueron cambiando conforme avanzó el proyecto, la

Tabla 9 señala el número de imágenes de la base de datos final para el entrenamiento de la IA para la detección de posiciones; como observación las imágenes para la fase de prueba fueron 100 imágenes que no forman parte de las imágenes para la fase de entrenamiento y validación. Cabe mencionar que el redimensionamiento final resulto en 200 x 200 pixeles en todas las imágenes y se usaron a color.

Tabla 9: Numero de imágenes totales para el entrenamiento de IA para la detección de posiciones.

Categoría	Imágenes iniciales	Imágenes a 90°	Imágenes a 180°	Imágenes a 270°	Total de imágenes	Imágenes para prueba
-----------	-----------------------	-------------------	--------------------	--------------------	----------------------	-------------------------

Normal	400	100	100	100	700	100
Lateral	400	100	100	100	700	100
Tapado	400	100	100	100	700	100
Boca abajo	400	100	100	100	700	100
Cabeza tapada	400	100	100	100	700	100
Total	2,000	500	500	500	3,500	500

Nota: El total de imágenes resulta de la suma de las imágenes iniciales sumando con las imágenes rotadas en los 3 ángulos señalados, las imágenes seleccionadas para prueba son 100 que no fueron utilizadas en las 2 fases previas (diseño propio).

De estas imágenes el quinto programa creado seleccionó 2,800 imágenes para la fase de entrenamiento (80%, 560 de cada categoría) y 700 imágenes para validación (20%, 140 de cada categoría). Las imágenes seleccionadas para fase de prueba exclusivamente se usaron para su fase.

Para la fase de prueba del programa de detección de posiciones con imágenes del bebé se realizó una selección de imágenes más representativas con diferentes características, como diferente iluminación, posición y ubicación, esto se realizó en cada categoría del total de imágenes tomadas al bebé seleccionando 100 de cada categoría tal como se señala en la

Tabla 10.

Tabla 10: Imágenes seleccionadas para la fase de prueba de bebé real.

Categoría	Imágenes iniciales tomadas al bebé	Imágenes seleccionadas
Normal	727	100
Lateral	396	100
Tapado	357	100
Boca abajo	465	100
ASSB	974	100
Total	2,919	500

Nota: A pesar de tener más imágenes para la detección de posiciones en cada categoría se seleccionaron las más representativas (principalmente no repetitivas y que tuvieran nitidez) y diversas iluminaciones. Se decidió para la fase de prueba que fuera un número en común de cada categoría, seleccionando 100 de cada una (diseño propio).

En el programa de detección de puntos rojos se creó una base de datos con 500 fotografías del maniquí para la fase de entrenamiento y se tomaron 50 fotografías del video del bebé durmiendo para la fase de prueba, obteniendo un total de 550 imágenes en total. Algunas características se detallan en la sección

Etapas 7: Creación de Programa para la Detección de Respiraciones.

Estas imágenes fueron realizadas con diferentes iluminaciones y ubicaciones del maniquí para su mejor detección; la

Figura 8 aparecen 4 ejemplos de fotografías tomadas con diversas iluminaciones y posiciones del maniquí para una mejor detección de puntos rojos.

Figura 8 Ejemplo de diferentes iluminaciones para la detección de puntos rojos.



Nota: Se puede apreciar principalmente diferencias en iluminaciones de las imágenes utilizadas para el entrenamiento del programa de detección de puntos rojos (fuente propia).

Análisis de Programa para la Detección de Posiciones

En esta sección hablaremos de los resultados obtenidos en la detección de posiciones donde el primer y segundo nivel de análisis se realizaron dentro de las fases de entrenamiento y validación mientras que el tercer nivel de análisis se realizó en 2 fases de prueba, una con maniquí y otra con imágenes de bebé.

Primer Nivel de Análisis: Imágenes de Maniquí

Después de varios intentos y pruebas el modelo final logro obtener resultados muy favorecedores después de la época 15. A partir de la época número 21 los resultados fueron muy buenos con respecto a la detección de posiciones en las imágenes obteniendo un 85% de *accuracy* en fase de entrenamiento. La pérdida en la fase de entrenamiento bajo continuamente hasta llegar al 28%.

En la fase de validación el *accuracy* supero en la época 13 el 85% esperado en esta investigación, logrando incluso mejorar en las siguientes épocas, mientras que *loss* (pérdida) bajo hasta el 3% en la época 23 señalando que el programa detecta muy bien las posiciones en la fase de validación en imágenes de maniquí. En la Tabla 11 se plasman los resultados obtenidos en la fase de entrenamiento y validación. Se presentan las épocas impares para no crear una tabla bastante grande.

Tabla 11: Resultados obtenidos de la fase de entrenamiento y validación del programa para la detección de posiciones.

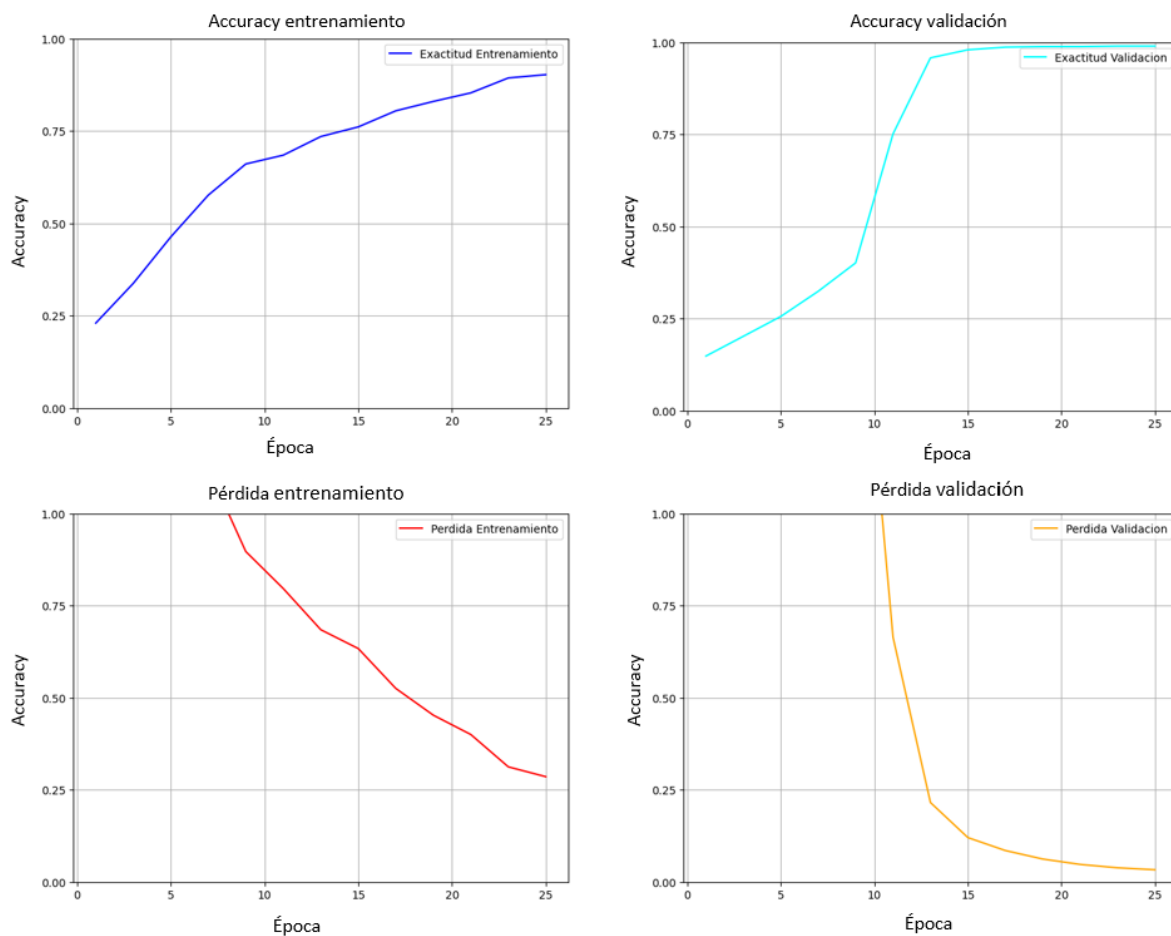
Época	Loss Entrenamiento	Accuracy Entrenamiento	Loss Validación	Accuracy Validación
1	2.4435	0.2300	1.8205	0.1483
3	1.9170	0.3375	2.6358	0.2020
5	1.4381	0.4632	2.1851	0.2558
7	1.1211	0.5764	2.4979	0.3241
9	0.8969	0.6607	1.7923	0.4012
11	0.7964	0.6846	0.6641	0.7515
13	0.6847	0.7350	0.2156	0.9578

15	0.6335	0.7611	0.1204	0.9797
17	0.5253	0.8046	0.0853	0.9869
19	0.4527	0.8300	0.0622	0.9884
21	0.4002	0.8532	0.0476	0.9884
23	0.3124	0.8936	0.0384	0.9898
25	0.2857	0.9025	0.0331	0.9898

Nota: A partir de la época 21 de la fase de entrenamiento se logra obtener poco más del 85% de *accuracy* esperado en esta investigación, mientras que en la fase de validación se logró superar el 85% a partir de la época 13, mejorando ambos hasta el 90% de *accuracy* en fase de entrenamiento y 98% en la fase de validación (diseño propio).

En las gráficas de la Figura 9 se puede observar de manera visual los resultados obtenidos en la fase de entrenamiento y validación, así como su mejora. Cabe mencionar que estas graficas se lograron creando otro programa dentro de *Colab* graficando los resultados.

Figura 9: Gráficas de resultados en fase de entrenamiento y validación.



Nota: Se puede observar la mejora de las métricas (*accuracy* y *loss*) mostrando con el paso de las épocas mayor *accuracy* y disminución de *loss* en cada fase (diseño propio).

Segundo Nivel de Análisis: Imágenes de Maniquí

La Tabla 12 compara los resultados obtenidos en la fase de entrenamiento y la validación resultando muy satisfactorios corroborando el funcionamiento del modelo final obtenido en la última época en fase de la validación.

Tabla 12: Comparación entre resultados finales obtenidos durante la fase de entrenamiento y validación.

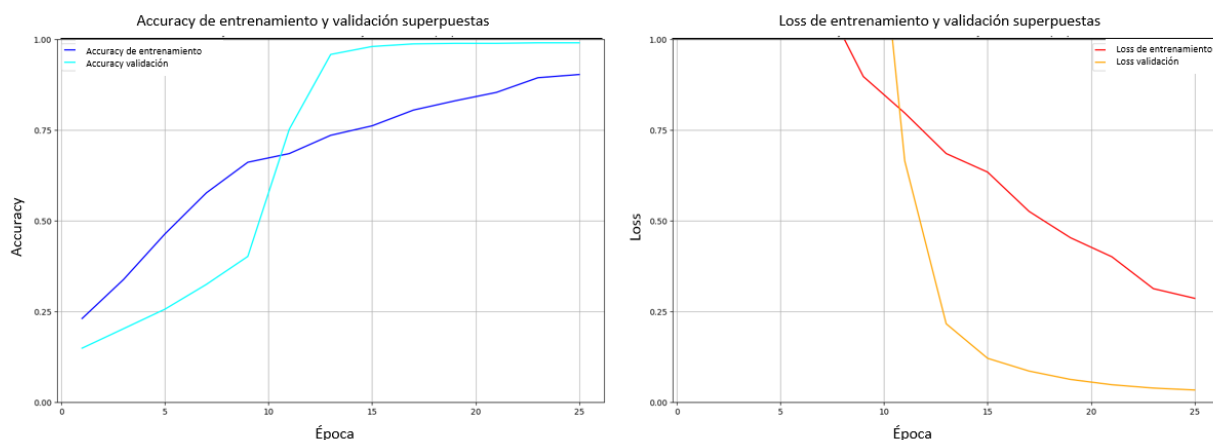
Ítem	Base de datos de maniquí	
	Fase	Resultado (%)

Accuracy	Entrenamiento	90.2
	Validación	98.9
Loss	Entrenamiento	28.5
	Validación	3.3

Nota: Debido a que se obtuvo en ambas fases un *accuracy* mayor al 85% esperado en esta investigación, se puede intuir que el modelo final puede ser utilizado para futuros proyectos (diseño propio).

La Figura 10 es una representación gráfica de las diferencias de las métricas de las 2 fases. Con el paso de las épocas fueron obteniendo mejores resultados a partir de la época 20 en cada uno siendo la fase de validación la que obtuvo los mejores resultados.

Figura 10: Gráficas superpuestas de accuracy y loss de las fases de entrenamiento y validación.



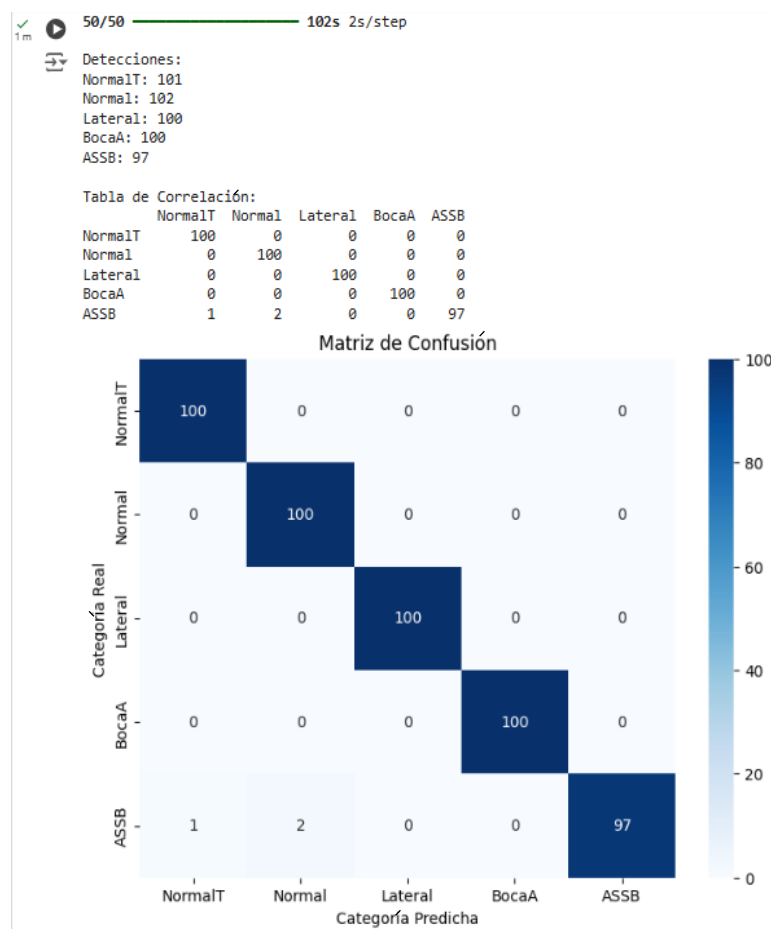
Nota: Los valores de la fase de validación mejoran a valores esperados antes que los valores de la fase de entrenamiento, por lo que el modelo aprende bien a seleccionar características específicas (fuente propia).

Segundo Nivel de Análisis: Correlación con Maniquí

Se realizó una fase de prueba con imágenes de maniquí que no se habían utilizado en la fase de entrenamiento ni validación. En esta fase de prueba se creó una tabla de correlación y una matriz de confusión, la Figura 11 es una captura de pantalla de los resultados obtenidos en el

programa. Se obtuvo un *Accuracy* de 99.4% y *loss* de 0.006 utilizando el modelo final, únicamente la categoría ASSB fue la que no se clasificó bien en tres ocasiones de 100, confundiendo 2 imágenes en la categoría Normal y una imagen en NormalT.

Figura 11: Resultados obtenidos en la fase de validación.



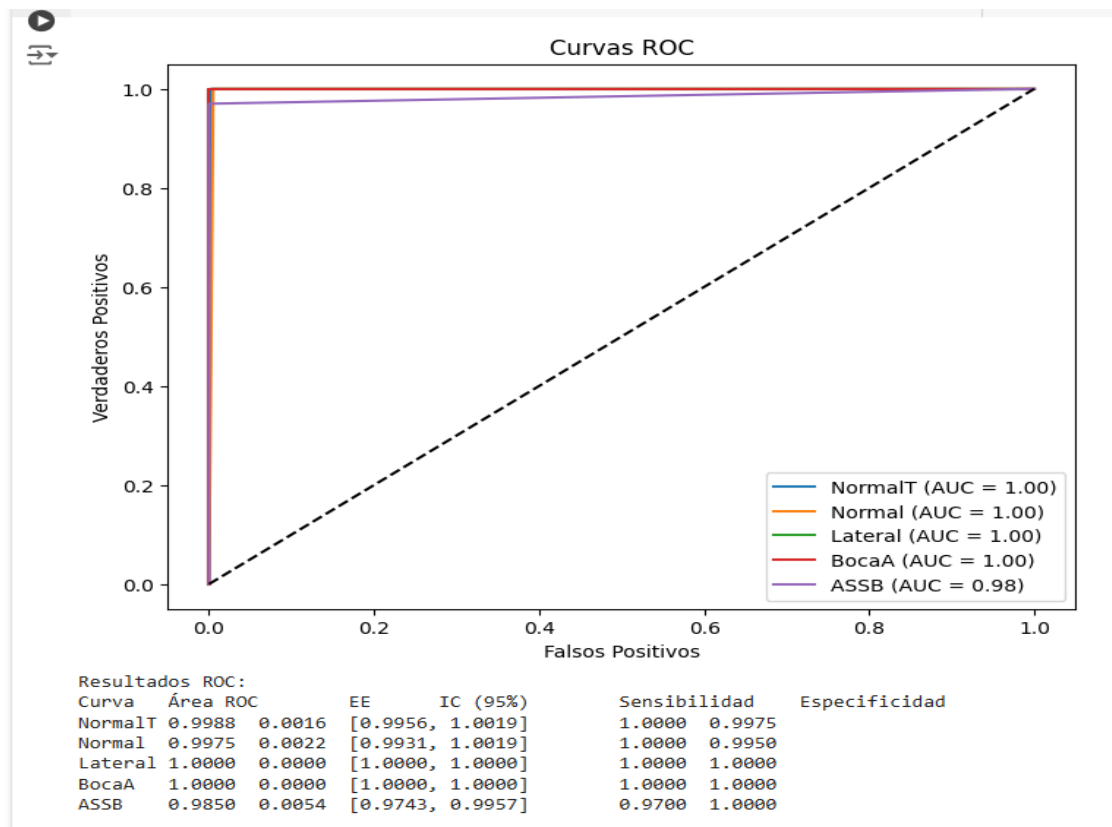
Nota: Ocurrieron 3 fallos de 100 en la detección de la categoría ASSB, en la cual una la considero como Normal y 2 como una posición Normal, todos los demás los detectó bien (fuente propia).

Tercer Nivel de Análisis: Imágenes de Maniquí

La Figura 12 muestra los resultados de obtenidos del tercer nivel de análisis, una gráfica representativa de la matriz de falsos positivos y negativos obtenidos en el segundo nivel de análisis. Se puede observar que el Área Debajo de la Curva (AUC por sus siglas en inglés) de todas las categorías (a excepción de ASSB) tienen un valor de 1 (100%), mientras que ASSB tiene un valor

de 0.97. Como es de esperarse en las categorías Lateral y BocaA el Error Estándar (EE) es de 0, debido que no existieron confusiones con ninguna otra categoría, demostrando que fueron las que tuvieron mejor resultado. En las 3 categorías faltantes el EE es muy bajo lo que indica una alta precisión, también se optó por obtener la sensibilidad (*Recall* o tasa de verdaderos positivos) y la especificidad (tasa de verdaderos negativos).

Figura 12: Curvas ROC de cada categoría y métricas de fase de prueba con fotos de maniquí.



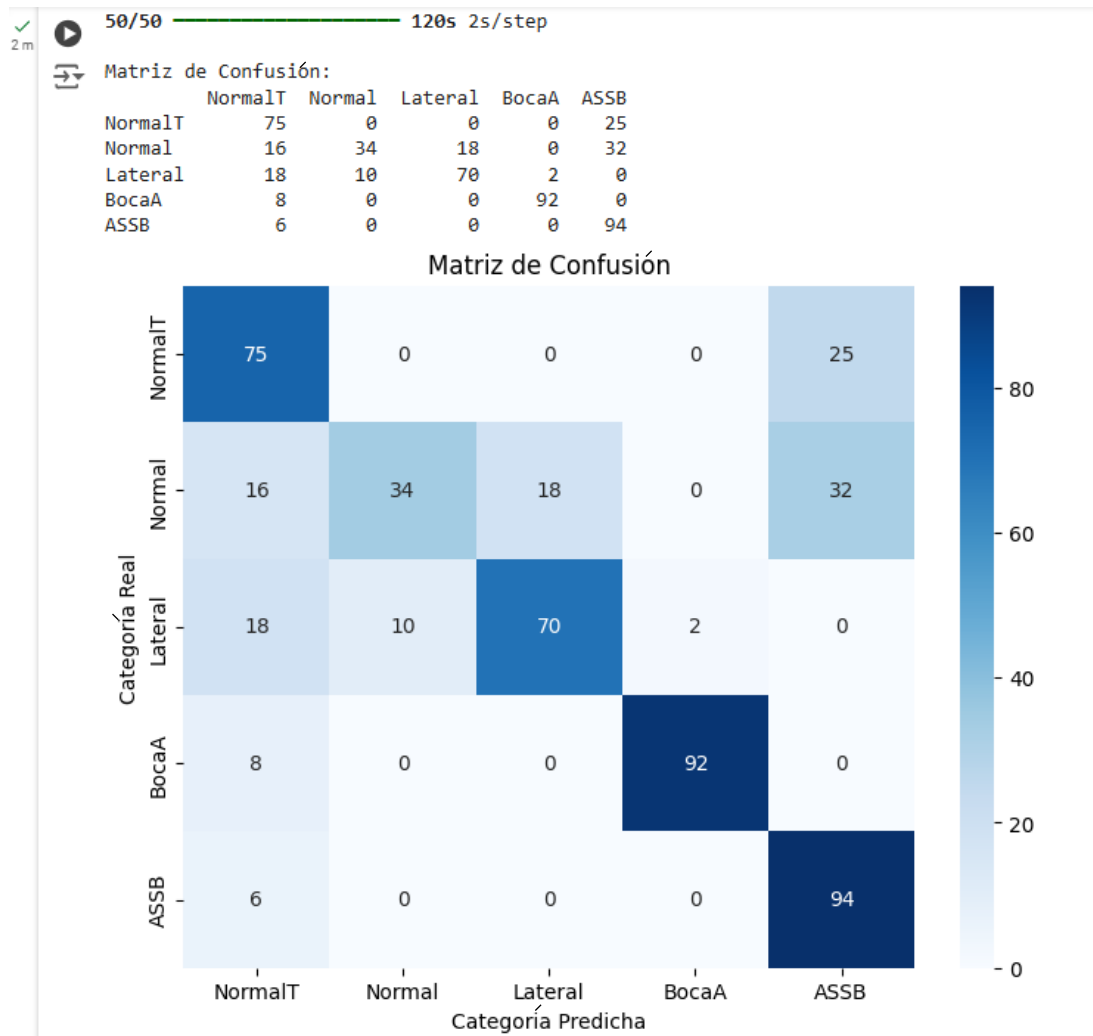
Nota: Las categorías Lateral y BocaA fueron las que obtuvieron un 100% de *accuracy* y no tuvieron confusión con otras categorías, siendo las que mejor clasifican imágenes (fuente propia).

Segundo Nivel de Análisis: Imágenes de Bebé

Como se mencionó en la metodología seleccionamos 500 imágenes del bebé en diferentes posiciones, 100 de cada categoría. Se realizó la fase de prueba con estas imágenes utilizando el modelo final obtenido de la fase de entrenamiento y se obtuvo una tabla de correlación tal como se muestra en la captura de pantalla de la Figura 13. Se obtuvo un *accuracy* de 0.73 y un *loss* de

0.27, las categorías que superaron el 85% esperado en esta investigación fueron ASSB y BocaA, mientras que la categoría Normal fue la que menos precisión obtuvo.

Figura 13: Resultados obtenidos de la fase de prueba con imágenes de bebé.



Nota: La categoría Normal fue la que tuvo menor resultado teniendo mayor confusión con ASSB, seguido de Lateral y NormalT demostrando que hace falta mejorar base de datos con imágenes en esas categorías. Solo 2 categorías (ASSB y BocaA) lograron el objetivo en esta investigación, señalando que se logró un buen entrenamiento en esas categorías (fuente propia).

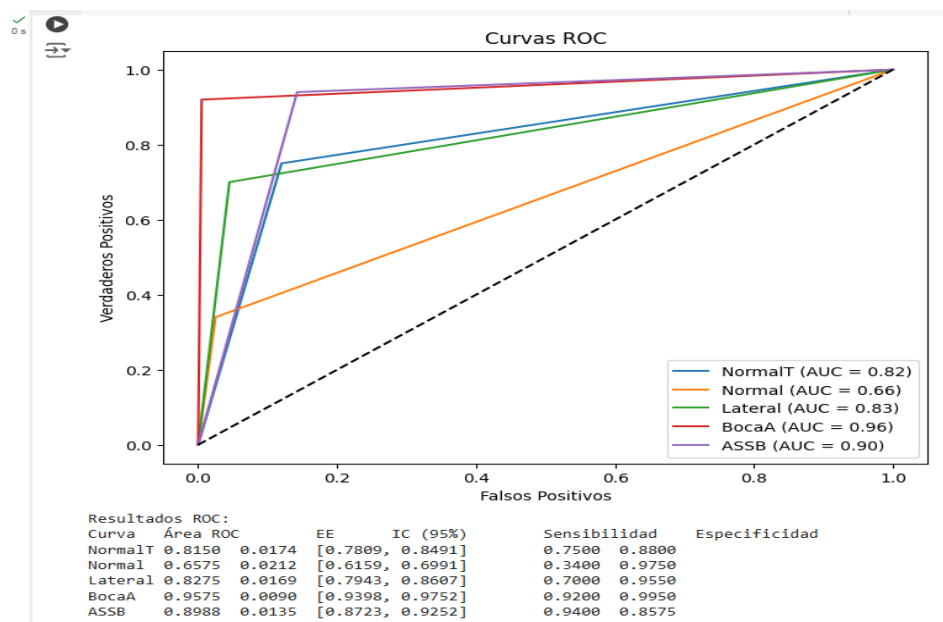
Existe una diferencia significativa entre los resultados obtenidos entre la base de datos con un maniquí y con la base de datos del bebé. Con la base de datos de maniquí se obtuvo un 99.4% de *accuracy*, muy superior al esperado en esta investigación (>85%), por otro lado, se obtuvo

solamente un *accuracy* del 73% con la base de datos del bebé. La categoría que mejor detecto fue ASSB demostrando que el modelo entrenó muy bien esta categoría y detecta objetos y mantas en la cara, pero la posición normal y lateral fueron las menos detectadas por lo que se necesita mejorar la base de datos en estas categorías.

Tercer Nivel de Análisis: Imágenes de Bebé

En base a la matriz obtenida en el segundo nivel de análisis se realizó un tercer nivel de análisis obteniendo las curvas ROC mostradas en la captura de pantalla de la Figura 14. Las categorías ASSB y BocaA obtuvieron un mayor AUC y menor EE, mientras que la categoría Normal fue la más afectada con menor AUC.

Figura 14: Curvas ROC de cada categoría y métricas de fase de prueba con fotos de bebé.



Nota: La categoría Normal tiene menor AUC y mayor EE, seguido de las categorías NormalT y Lateral, las categorías restantes muestran mejores resultados (fuente propia).

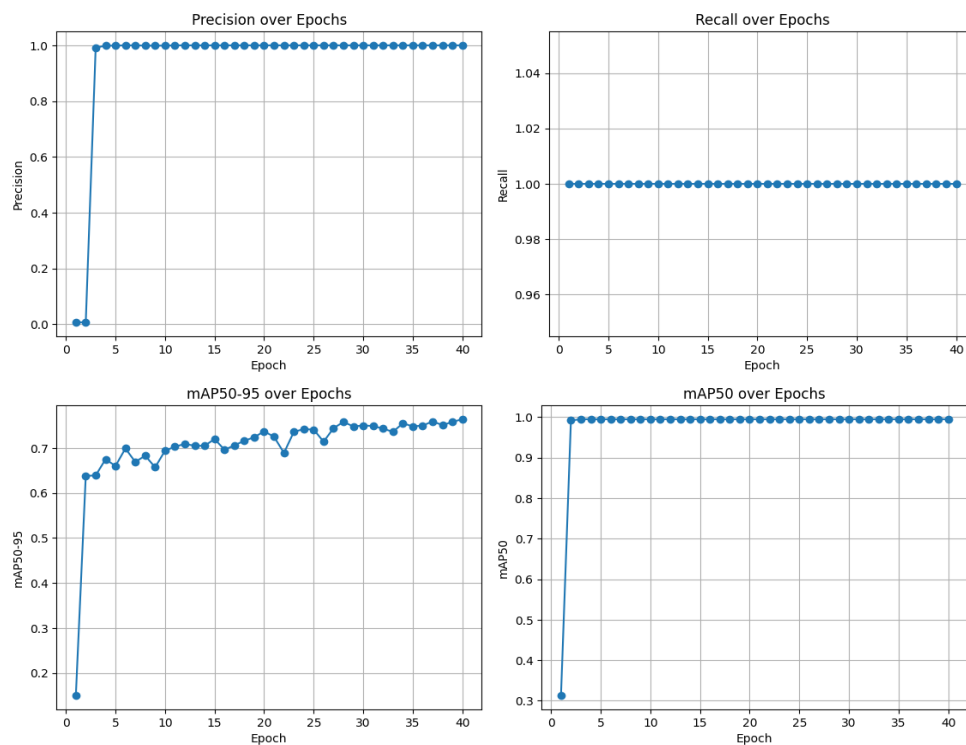
Análisis de Programa para la Detección de Respiraciones

Como se mencionó en la metodología para la detección de respiraciones se creó una IA basada en YOLOv8 por lo que se analizaron los resultados de los tres niveles de análisis en la detección de los puntos rojos en el abdomen.

Primer Nivel de Análisis

Después de realizar las 40 épocas se obtuvieron resultados favorecedores con respecto a la detección de los puntos rojos, la Figura 15 se conforma de las 4 graficas de los resultados obtenidos de la fase de entrenamiento final, a partir de la época número cinco se obtuvo una estabilidad en *accuracy*, *Recall* y mAP50, incluso al 100% de efectividad, mientras que mAP50-95 no lo obtuvo, pero siguió con una tendencia a mejorar.

Figura 15: Resultados de la fase de entrenamiento de YOLOv8 para la detección de puntos rojos.



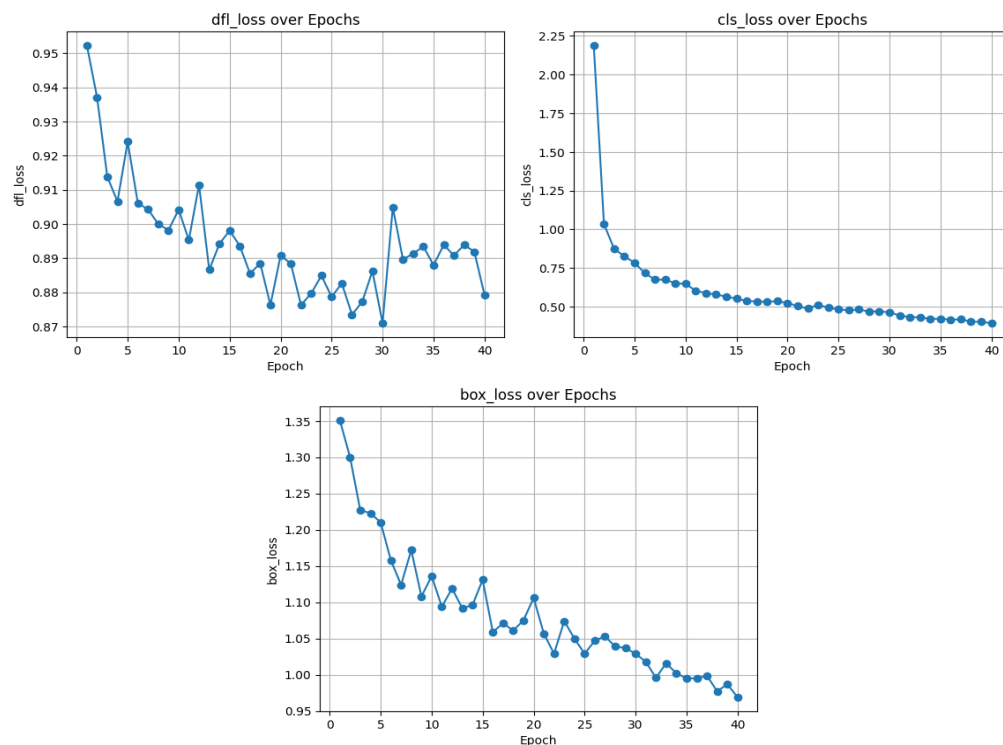
Nota: Los resultados de la fase de entrenamiento fueron muy favorecedores desde las primeras épocas, demostrando una alta eficacia (fuente propia).

Una característica de YOLO es que maneja 3 métricas de *loss*: *Dfl_loss* es la predicción de la distribución de los objetos a detectar, esto es que el modelo intenta predecir la ubicación de los puntos rojos en cada una de las imágenes para saber dónde pueden estar y detectarlos más fácilmente, *Cls_loss* es la pérdida de clasificación señalando que tan bien o mal detecta la

clasificación del objeto a detectar, y *Box_loss* mide que tan precisa es la ubicación y el tamaño de las cajas delimitadoras.

En la Figura 16 se muestran los resultados obtenidos de los diferentes *loss*. A pesar no obtener un valor por debajo de 10% en los diferentes tipos de error se detectaban muy bien los puntos rojos durante la evaluación de imagen. Al observar estos resultados satisfactorios se obtuvo el mejor modelo de procesamiento y fue utilizado en la fase de validación y de prueba.

Figura 16: Resultados de diferentes tipos de *loss* en fase de entrenamiento de Yolov8.



Nota: Se puede observar el descenso rápido de los diferentes *Loss*, principalmente *cls loss* desde las primeras épocas debido a que solo era una clasificación, debido también a que se utilizaron diversas ubicaciones del maniquí en la cuna *dfl_loss* tuvo mayor variación (fuente propia).

Segundo Nivel de Análisis

En la fase de validación se obtuvieron resultados favorables. Los resultados tanto de la fase de entrenamiento como de validación se colocaron en la Tabla 13, los resultados de la época 40 de la fase de entrenamiento y la fase de validación fueron resultados casi idénticos mostrando que el



modelo final entrenó de manera satisfactoria obteniendo un 100% de reconocimiento de los 2 puntos rojos en el abdomen.

Tabla 13: Resultados de fase de validación y comparación con los resultados de la última época de la fase de entrenamiento.

Métrica	Época 40 (%)	Validación (%)
Precisión	100	100
<i>Recall</i>	100	100
mAP50	99.5	99.5
mAP50-95	75.5	75.7

Nota: Resultados satisfactorios superando el 85% esperado en esta investigación demostrando ser un modelo final muy efectivo en la detección de los puntos rojos en el abdomen (diseño propio).

Para realizar la fase de prueba se utilizaron las 50 imágenes seleccionadas para esta fase en *Roboflow*, obteniendo una efectividad del 100% y al obtener este porcentaje no se pudo obtener una matriz de confusión.

Al mismo tiempo se intentó hacer una tabla de correlación, sin embargo, al no haber variaciones en los resultados el programa entregó “NaN” en cada cuadro; con esto se infiere que el programa logra detectar el 100% de los puntos rojos.

La Figura 17 es una captura de pantalla donde se muestran los resultados obtenidos en el programa demostrando lo antes mencionado; la clasificación “No detectado” aplica si el programa hubiera detectado solo 1 punto rojo, ninguno o más de 2 puntos rojos.

Figura 17: Resultados de la fase de prueba del programa Yolo.

Reporte clasificacion:				
	precision	recall	f1-score	support
No detectado	0.00	0.00	0.00	0
Detectado	1.00	1.00	1.00	50
accuracy			1.00	50
macro avg	0.50	0.50	0.50	50
weighted avg	1.00	1.00	1.00	50
Matriz confusion:				
[[50]]				
Tabla de correlacion:				
	Imágenes	Detecciones		
Imágenes	NaN	NaN		
Detecciones	NaN	NaN		

Nota: Obtención de resultados con otro programa en *Colab* utilizando el mejor modelo obtenido en la fase de entrenamiento y las imágenes seleccionadas para esta fase. La variable “No detectado” corresponde a no haber detectado puntos o solo un punto rojo en la imagen, mientras que “Detectado” es que detectaba los 2 puntos rojos en cada imagen (fuente propia).

Tercer Nivel de Análisis

Se buscó encontrar regresión logística con los resultados obtenidos, pero al tener el 100% de lecturas no se puede obtener una regresión; se realizó otra fase de prueba con las imágenes del bebé obteniendo el mismo resultado. Por lo tanto, se optó por realizar otra fase de prueba con 100 imágenes del programa para la detección de posiciones, 60 imágenes fueron seleccionadas al azar de la categoría Normal con el maniquí en diversas posiciones e iluminaciones y 40 imágenes de la base de datos del bebé utilizadas en el mismo programa, cabe aclarar que estas imágenes nunca las vio el programa de detección de puntos rojos por lo que eran ideales para la fase de prueba.

Al realizar la prueba con esas 100 imágenes nuevas se obtuvieron muy buenos resultados, a diferencia de la fase de prueba anterior en tres imágenes no se detectó correctamente dos puntos (en dos imágenes se detectaron tres puntos y en una imagen se detectó solo un punto), sin embargo, el *accuracy* resulto ser del 100% debido a que las imágenes que predijo que tenía solo dos puntos fueron efectivamente detectados (verdaderos positivos).

Se realizó una matriz donde se observó que la clase “Detectados” (fila inferior) contiene también la detección de los tres verdaderos negativos además de los 97 verdaderos positivos, por lo que la clase “No detectados” (fila superior) no tiene ningún valor y no se puede realizar una regresión logística. La Figura 18 es una captura de pantalla una vez terminada la fase de prueba, estos resultados señalan que el modelo final es muy eficiente al momento de detectar tanto los puntos rojos como cuando no se detecta correctamente los 2 puntos rojos.

Figura 18: Resultados obtenidos en fase de prueba con 100 imágenes aleatorias.

Reporte clasificacion:				
	precision	recall	f1-score	support
No detectado	0.00	0.00	0.00	0
Detectado	1.00	0.97	0.98	100
accuracy			0.97	100
macro avg	0.50	0.48	0.49	100
weighted avg	1.00	0.97	0.98	100
Matriz confusion:				
[[0 0]				
[3 97]]				

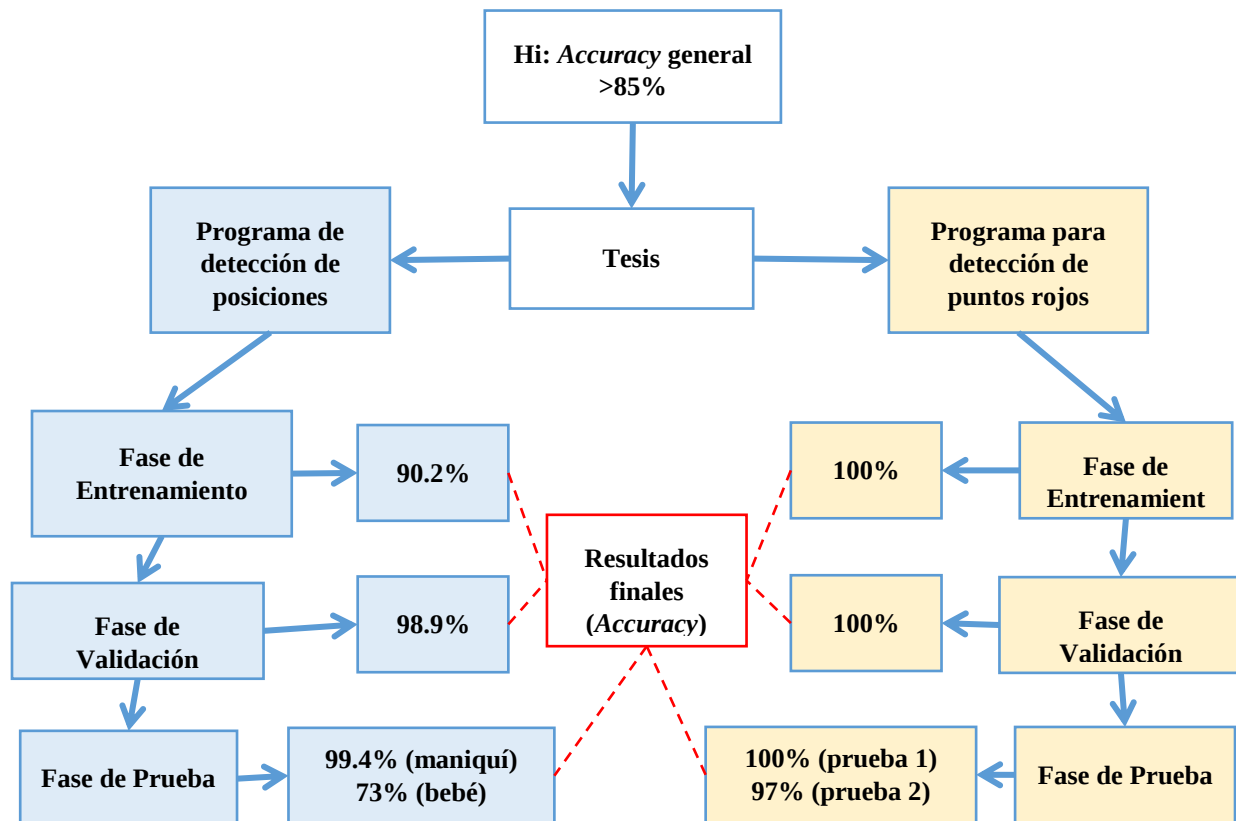
Nota: El *accuracy* demostró ser de 100% logrando detectar correctamente los verdaderos positivos, además el modelo supo diferenciar cuando no se detectaban correctamente los 2 puntos rojos en 3 imágenes, dando un *Recall* de 97% (fuente propia).

Tesis

Se obtuvieron 2 modelos entrenados que lograron en las fases de prueba con el maniquí un *accuracy* del 99.4% en el programa de detección de posiciones y 100% en detección de puntos rojos. En la fase de prueba con las imágenes del bebé real se obtuvieron menores resultados (73% de *accuracy*) en el programa de detección de posiciones, pero el 100% en la detección de puntos

rojos. Para aumentar la precisión de los modelos con imágenes de bebés reales se necesita entrenar los modelos obtenidos con imágenes de bebés con diversas características como diferentes tamaños, tonos de piel y edades. Observando el buen desempeño de los modelos con las imágenes del maniquí se espera que sea el mismo caso cuando se tenga una base de datos más variable de bebés para poder ser utilizados en el futuro para el diseño de un sistema integra. La Figura 19 es un diagrama de tesis de cómo se dieron los resultados.

Figura 19: Diagrama de tesis.



Nota: Se inicio con una hipótesis, después se planteó la metodología y lo que se obtendría, por último, se muestran los resultados de cada una de las fases (diseño propio).

DISCUSIÓN

Esta investigación resalta la obtención de más del 97% de *accuracy* en fases de prueba con maniquí en ambos programas de los que en el programa de detección de posiciones se obtuvo más del 90% de *accuracy* en las fases de entrenamiento y validación (mayor al 85% esperado en esta

investigación) mientras que en la fase de prueba con imágenes del maniquí se obtuvo un *accuracy* del 99.4%, sin embargo, en la fase de prueba con las imágenes de bebé real se obtuvo un *accuracy* de 73%. Esta disminución se puede atribuir a variables como la diferencia en el color de piel, colores del ambiente de prueba, texturas, entre otros. Estas son áreas de oportunidad para el reentrenamiento del modelo.

Con estos resultados se puede inferir que este modelo final es funcional y se requiere entrenar el modelo con una base de datos mayor de bebés reales para obtener un *accuracy* mejor en la detección de posiciones en un entorno real; el programa de detección de puntos rojos logró alcanzar el 100% de *accuracy*, más de lo esperado como objetivo de esta investigación. En la tercera prueba con imágenes tomadas del proyecto de detección de posiciones se obtuvo un *accuracy* del 97% y a diferencia del programa anterior, este modelo final puede ser implementado para la detección de puntos rojo y así medir la respiración.

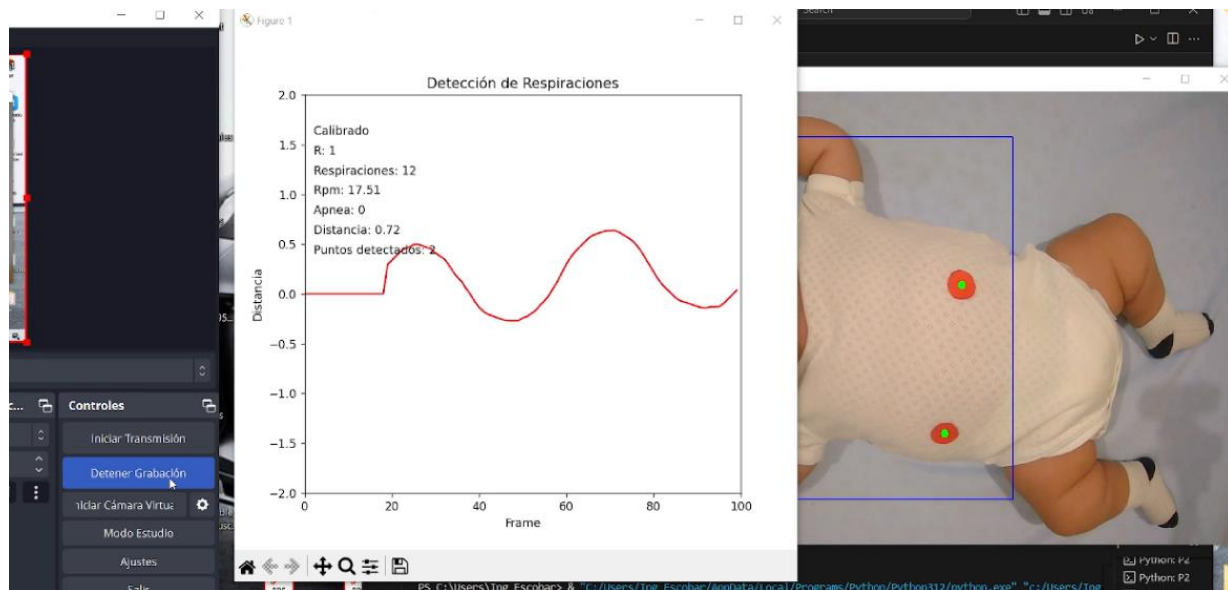
Zhu et al. [14] realizaron un programa para la detección de objetos que promuevan una muerte de cuna pero obtuvieron un 80% en la fase de prueba a diferencia de esta investigación que se obtuvo más del 95% de detección de objetos. Adewopo et al. [10] trabajaron 4 modelos de algoritmos para la detección de movimientos y acciones del bebé obteniendo un *accuracy* entre el 80 y 99%, en esta investigación se logró obtener resultados parecidos en la fase de prueba con imágenes del maniquí.

En la investigación de Joseph C. Afonne [12] implemento el modelo YOLO obteniendo un *accuracy* mayor a 97% señalando que los programas creados con este modelo se obtienen muy buenos resultados en la detección de objetos. A diferencia de la investigación realizada por Bharati [11] que solamente trabajó con 3 categorías (boca arriba, transición a boca abajo y boca abajo) en esta investigación se trabajaron 5 categorías aumentando el rango de factores de riesgo en un mismo programa.

Se pueden desarrollar 2 programas complementarios que utilicen los modelos creados en esta investigación para detectar factores de riesgo de muerte de cuna con una sola cámara si se emplean en un mismo dispositivo.

Como ejemplo, se desarrolló un programa complementario donde se emplea el modelo final de detección de puntos rojos utilizando el video tomado del bebé respirando. La Figura 20 es una captura de pantalla ejecutando el programa complementario donde se puede observar cómo detecta los puntos rojos y puede medir tanto frecuencia respiratoria como apneas.

Figura 20: Captura de pantalla del programa complementario para demostrar funcionamiento de programa de detección de puntos rojos.



Nota: Durante la ejecución del programa se pueden medir las respiraciones por minuto, la cantidad de puntos detectados y cuantas respiraciones lleva desde que inició el programa, cuando no detecta respiraciones o movimientos de los puntos por más de 5 segundos comienza un conteo de apnea a partir del número 5 señalando los segundos sin respirar (fuente propia).

Mientras que la mayoría de las investigaciones encontradas detectan solo algunos factores de riesgo por separado mientras que en esta investigación se señala que se puede detectar objetos, posiciones, cuando se tapa la cara, cuando este tapado hasta el abdomen y también la detección de puntos rojos en el abdomen para la detección de respiraciones, dando un rango de detección de factores de riesgo de muerte de cuna más amplio. Una mejora para futuras investigaciones puede ser el uso de cámara nocturna para la obtención de bases de datos con escasa iluminación, así como



su implementación. Otra mejora sería el utilizar otra manta de color diferente para que los modelos entrenen con diferentes fondos y detecten el bebé en otros entornos.

Se puede crear una base de datos mayor con bebés de diferentes edades menores a un año, de diferente color de piel, tamaño, rasgos y diversos fondos de cunas con el fin de entrenar modelos más realistas y así mejorar la precisión en entornos reales.

También el set-up puede llegar a ser más ergonómico y de mejores materiales, en proyectos futuros se pueden crear más categorías y crear un sistema integral mejorado. Un ejemplo de otras categorías que se pueden integrar es detectar cuando existan manos manipulando al bebé y detección de sonido. Los modelos creados en esta investigación serán mejorados y utilizados en un futuro para la creación de un sistema integral.

CONCLUSIONES

Se logró obtener en todas las fases de los dos programas un *accuracy* mayor al esperado en esta investigación (mayor a 85%) con el maniquí. Contrastando los resultados obtenidos en la fase de prueba con imágenes del bebé, nos indica que es necesario crear una base de datos grande de varios bebés reales para entrenar los modelos y así obtener un mejor *accuracy* para un entorno real. Estos modelos permitirán en un futuro desarrollar programas que detecten con alta precisión las posiciones y las respiraciones de un bebé en tiempo real por medio de una cámara. Al obtener solamente la información de una cámara se puede desarrollar un equipo sencillo que no afecte al bebé y que sea económico.

BIBLIOGRAFÍA

1. Moon RY, Carlin RF, Hand I, Task Force on Sudden Infant Death Syndrome and the Committee on Fetus and Newborn. Sleep-Related Infant Deaths: Updated 2022 Recommendations for Reducing Infant Deaths in the Sleep Environment. *Pediatrics*. 2022;150:e2022057990. doi:10.1542/peds.2022-057990

2. Martínez-Valdez L, Richardson V, Bautista-Márquez A, Hernández-Ávila M. Epidemiology of sudden infant death syndrome in Mexico, 2005–2020. *Front Pediatr.* 2022;10. doi:10.3389/fped.2022.1001089
3. PAHO. Clasificación estadística internacional de enfermedades y problemas relacionados con la salud: Volumes 1, 2 & 3. 10th ed. Washington, D. C: Pan American Health Organization; 2002.
4. Vincent A, Chu NT, Shah A, Avanthika C, Jhaveri S, Singh K, et al. Sudden Infant Death Syndrome: Risk Factors and Newer Risk Reduction Strategies. *Cureus.* 2023;15:e40572. doi:10.7759/cureus.40572
5. Kaler J, Hussain A, Lee S. Manifestation and Pathogenesis of Sudden Infant Death Syndrome: A Review. *Crit Rev Eukaryot Gene Expr.* 2020;30. doi:10.1615/CritRevEukaryotGeneExpr.2020033009
6. Hasan MNU, Negulescu I. Wearable Technology for Baby Monitoring: A Review. *J Text Eng Fash Technol.* 2020;6:112-20. doi:10.15406/jteft.2020.06.00239
7. Lorato I, Stuijk S, Meftah M, Kommers D, Andriessen P, Pul C van, et al. Multi-camera infrared thermography for infant respiration monitoring. *Biomed Opt Express.* 2020;11:4848-61. doi:10.1364/BOE.397188
8. Reshmi R, Saranya RA, Vishakha S, Prabha DSU, Ashokan MS. Sudden infant death syndrome detection and prevention using grid eye sensor. *Int Res J Eng Technol.* 2019;06:623-5.
9. Kim JD, Lee WH, Lee Y, Lee HJ, Cha T, Kim SH, et al. Non-contact respiration monitoring using impulse radio ultrawideband radar in neonates. *R Soc Open Sci.* 2019;6:190149. doi:10.1098/rsos.190149
10. Adewopo V, Elsayed N, Anderson K. Baby Physical Safety Monitoring in Smart Home Using Action Recognition System. En: *SoutheastCon 2023.* 2023. p. 142-9. doi:10.1109/SoutheastCon51012.2023.10115146
11. Bharati VS. An Efficient Convolutional Neural Network Computer Vision System to Prevent Sudden Infant Death Syndrome. *Engineering Archive;* 2020. doi:10.31224/osf.io/wbx63
12. Joseph C Afonne. Towards Sustainable Oceans: Deep Learning Models for Accurate COTS Detection in Underwater Images [Tesis de maestría] [Internet]. [Finlandia]: University of Vaasa; 2024 [citado 29 de agosto de 2024]. Disponible en: <https://osuva.uwasa.fi/items/eb59c9c8-b636-45e2-a204-8130c71e37e5>
13. Klicker L, Maliwat A, Findura J, Yu X, Gromov M, Saniie J. InfaSafe: A Comprehensive, Non-invasive Infant Monitoring System. En: *2024 IEEE International Conference on Electro Information Technology (eIT).* Eau Claire, WI, USA: IEEE; 2024. p. 291-6. doi:10.1109/eIT60633.2024.10609853

14. Zhu S, Mathew A, Hatamimajoumerd E, Wan M, Taylor B, Venkatesaramani R, et al. CribNet: Enhancing Infant Safety in Cribs Through Vision-Based Hazard Detection. En: 2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG). Istanbul, Turkiye: IEEE; 2024. p. 01-8. doi:10.1109/FG59268.2024.10581871
15. Organización Mundial de la Salud. Nacimientos prematuros [Internet]. [citado 24 de agosto de 2024]. Disponible en: <https://www.who.int/es/news-room/fact-sheets/detail/preterm-birth>
16. Munkel Ramírez L, Durón González R, Bolaños Morera P. Síndrome de muerte súbita del lactante. Med Leg Costa Rica. 2018;35:65-74.
17. Otero Sanchez de Delgado LM. Desarrollo de una Cuna Interactiva para Monitoreo y Control del Bienestar del Bebé - 2023 [Tesis de Licenciatura]. [Perú]: Universidad Privada del Norte; 2023.
18. Chiu TKF, Ahmad Z, Ismailov M, Sanusi IT. What are artificial intelligence literacy and competency? A comprehensive framework to support them. Comput Educ Open. 2024;6:100171. doi:10.1016/j.caeo.2024.100171
19. Gumbs AA, Perretta S, d'Allemagne B, Chouillard E. What is Artificial Intelligence Surgery? Artif Intell Surg. 2021;1:1-10. doi:10.20517/ais.2021.01
20. Álvaro AM. Clasificación de imágenes usando redes neuronales convolucionales en Python [Tesis de Licenciatura]. [Sevilla]: Universidad de Sevilla; 2019.
21. Garbin C, Zhu X, Marques O. Dropout vs. batch normalization: an empirical study of their impact to deep learning. Multimed Tools Appl. 2020;79:12777-815. doi:10.1007/s11042-019-08453-9
22. Wolf JH. Rest Uneasy: Sudden Infant Death Syndrome in Twentieth-Century America by Brittany Cowgill (review). Bull Hist Med. 2019;93:290-1.
23. Alkharabsheh AR. An Intelligent Mobile Agents System for Sudden Infant Death Syndrome Monitoring. En: 2018 Advances in Science and Engineering Technology International Conferences (ASET). Abu Dhabi: IEEE; 2018. p. 1-7. doi:10.1109/ICASET.2018.8376888
24. Ho ESL, McCay KD, Marcroft C, Embleton ND. PCPP: A MATLAB application for abnormal infant movement detection from video. Softw Impacts. 2022;14:100412. doi:10.1016/j.simpa.2022.100412
25. Zúñiga PIV, Cedeño RJC, Palacios IAM. Metodología de la investigación científica: guía práctica. Cienc Lat Rev Científica Multidiscip. 2023;7:4. doi:10.37811/cl_rcm.v7i4.7658
26. Shahinfar S, Meek P, Falzon G. "How many images do I need?" Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. Ecol Inform. 2020;57:101085. doi:10.1016/j.ecoinf.2020.101085



27. Ballesteros Rincón JS, Guzmán Trillos OD. Programa informático de apoyo para la evaluación del desarrollo de tumores mediante el método RECIST aplicando procesamiento digital de imágenes biomédicas a través de inteligencia artificial [Tesis de Licenciatura] [Internet]. Universidad Autónoma de Bucaramanga; 2022 [citado 29 de agosto de 2024]. Disponible en: <https://repository.unab.edu.co/handle/20.500.12749/18965>
28. Polit DF, Beck CT. The content validity index: are you sure you know what's being reported? Critique and recommendations. *Res Nurs Health*. 2006;29:489-97. doi:10.1002/nur.20147
29. Yani M, Budhi Irawan S, Casi Setiningsih S. Application of Transfer Learning Using Convolutional Neural Network Method for Early Detection of Terry's Nail. *J Phys Conf Ser*. 2019;1201:012052. doi:10.1088/1742-6596/1201/1/012052

ANEXOS

Anexo 1. Formato de carta de consentimiento

Consentimiento informado parte 1.



UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA
FACULTAD DE MEDICINA Y CIENCIAS BIOMÉDICAS



Carta de consentimiento informado:

Fecha: ____ / ____ / ____
Chihuahua, Chihuahua

Título de proyecto: Detección de Factores de Riesgo de Muerte de Cuna por Medio de Inteligencia Artificial

Investigador: Ing. Alexis Salvador Escobar Beserra

Maestría: Formación e Innovación en Profesionales de la Salud

Institución: Facultad de Medicina de la Universidad Autónoma de Chihuahua

¿Cuál es el propósito de esta investigación?

Desarrollar 2 programas para la detección oportuna de factores de riesgo de muerte de cuna por medio de una cámara. El primer programa consiste en utilizar Inteligencia Artificial (IA) para la detección de las siguientes categorías (posiciones): bebé boca arriba, boca abajo, de lado, tapado y con cobija en cara u objetos. El segundo programa realizará la función de detección de respiraciones, detectando 2 puntos rojos ubicados en el abdomen de una prenda diseñada para un maniquí y el bebé. El propósito de este consentimiento es solicitar la autorización para tomar fotografías y usarlas para la corroboración de funcionamiento de estos programas.

Confidencialidad y privacidad

Toda la información recolectada es confidencial. Las imágenes y video serán tratadas como datos protegidos y únicamente el investigador tendrá acceso a ellas. El programa convertirá las imágenes en datos numéricos y en todo momento el rostro será pixelado para proteger la identidad del bebé. Una vez terminada la investigación dicha información recolectada será eliminada.

¿En qué consiste la participación?

Si acepta participar en esta investigación, realizaremos pruebas de observación con su bebé:

- Se colocará a su bebé en una cuna equipada con una cámara diseñada por el investigador
- Se realizará la toma de fotografías en las posiciones que se señalan en la *Tabla 1*
- Se tomará un video de 30 segundos del bebé en posición boca arriba para la detección de respiraciones
- Tiempo: las pruebas son breves y no interfieren con el cuidado y/o alimentación del bebé

Tabla 1

Posición	Acción
Boca abajo	Posición de mayor riesgo para el bebé. Se considera que esta posición es 3 veces más propensa a una muerte de cuna.
Boca arriba	Es la posición más segura para que un bebé duerma. Se deben detectar los puntos rojos del abdomen y no se debe detectar mantas u objetos dentro del área de cuna.
Lateral	Se considera una posición de riesgo para muerte de cuna. Puede ser con un objeto en la espalda que mantenga al bebé en posición lateral. Se puede llegar a ver solo 1 punto rojo del abdomen.

Consentimiento informado parte 2.



UNIVERSIDAD AUTÓNOMA DE CHIHUAHUA FACULTAD DE MEDICINA Y CIENCIAS BIOMÉDICAS



Con cobija	No se considera un factor de riesgo, esta posición es cuando se encuentra boca arriba o lateral con una manta que tape sus piernas y abdomen sin sobrepasar el nivel inferior de los pezones.
Cobija tapando cabeza o detección de objetos	Se les considera también un factor de riesgo para muerte por asfixia. En este tipo se colocarán objetos alrededor del bebé sin acercarlos demasiado ni incomodar al bebé. En cuestión de la manta se sobrepondrá al bebé sin molestarlo simulando que tiene una manta cerca de la cara.

Riesgos

La investigación se clasifica como riesgo mínimo, no se aplicará ningún procedimiento invasivo. La cámara y los materiales con las que esta hecha la cuna son seguros y no tiene contacto directo con la piel. La prenda que se utilizará será nueva y lavada para evitar cualquier alérgeno.

Declaración del padre o tutor

He leído la información anterior y he tenido la oportunidad de hacer preguntas y todas han sido contestadas. Comprendo que la participación es voluntaria y que puedo retirar a mi hijo/a de esta investigación en cualquier momento sin algún perjuicio a nosotros. Al firmar este documento doy mi consentimiento para:

- ☐ Autorizo la participación de mi hijo/a en esta investigación, conociendo el alcance, las actividades y riesgos que conlleve para el bebé
- ☐ Autorizo el procesamiento de imágenes y video exclusivamente para los fines de esta investigación
- ☐ Entiendo que los datos serán tratados de manera anónima y segura
- ☐ Aclaramos que esta participación es completamente voluntaria
- ☐ Autorizo el uso de una fotografía como evidencia del proyecto y únicamente para fines de esta investigación, cuidando el anonimato del bebé

Contacto:

Para cualquier duda, aclaración o pregunta se puede contactar por los siguientes medios:

Investigador: Ing. Alexis S. Escobar Beserra.

Directora de tesis: Dra. Karina Rodríguez Carmona

Facultad de Medicina y Ciencias biomédicas, Universidad Autónoma de Chihuahua. Tel: 614-238-6030

Nombre y firma de padre o tutor

Firma de padre o tutor

Nombre y firma del investigador

Anexo 2. Aprobación de Protocolo de Investigación



UNIVERSIDAD AUTÓNOMA DE
CHIHUAHUA

Chihuahua, Chih., 11 de noviembre de 2025
Asunto: Dictamen de protocolo

**DRA. KARINA RODRÍGUEZ CARMONA
PRESENTE.-**

Por este medio notificamos a usted que, el Comité de Ética en Investigación de la Facultad de Medicina y Ciencias Biomédicas, revisó el protocolo titulado ***"DETECCIÓN DE FACTORES DE RIESGO DE MUERTE DE CUNA POR MEDIO DE INTELIGENCIA ARTIFICIAL"***, con número de registro **CI-030-23** y se dictaminó como **APROBADO**.

Sin más por el momento, quedo de Usted.

ATENTAMENTE
"MENTI DA LUCEM, MANIBUS ARTEM"


DR. JAIR CARRACO PALAFOX
PRESIDENTE DEL COMITÉ DE ÉTICA
EN INVESTIGACIÓN FMycB



FACULTAD DE MEDICINA Y CIENCIAS BIOMÉDICAS
SECRETARÍA DE INVESTIGACIÓN Y POSGRADO
Circuito Universitario, Campus II C.P. 31125
Teléfonos: 52(614) 238.6030 / (614) 439.3500 ext. 3534
Chihuahua, Chih., Mex.
www.fm.uach.mx

