



Universidad Modelo

Escuela de Ingenierías – Ingeniería Biomédica

Protocolo de investigación

Título: Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla

Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla

Chávez Gómez Jorge Alejandro

Monterrubio Antonio María Fernanda

Abstract

Background: Clinical quantification of dynamic knee valgus (a key risk factor for ACL tears in athletes) is restricted by visual subjectivity due to high-speed impact. Objective: To develop a computer vision system for the automated identification and classification of this risk. Methodology: Frontal video recordings (60 fps) of the Single Leg Drop Jump test (a 30-cm box drop followed by a maximal single-leg vertical jump) were analyzed using a hybrid network (CNN ResNet50 + LSTM) to isolate the frame of maximum articular collapse. Subsequently, MediaPipe BlazePose extracted 3D hip, knee, and ankle coordinates to calculate the Frontal Plane Projection Angle (FPPA) via vector geometry for injury risk classification.

Resumen

Antecedentes: La cuantificación del valgo dinámico de rodilla (riesgo clave para la ruptura del LCA en atletas) se ve limitada en la clínica por la subjetividad visual ante la velocidad del impacto. Objetivo: Desarrollar un sistema de visión artificial para la identificación y clasificación automática de este riesgo. Metodología: Se analizaron videos frontales (60 fps) de la prueba Single Leg Drop Jump (caída desde 30 cm y salto monopodal máximo) mediante una red híbrida (CNN ResNet50 + LSTM) para aislar el fotograma de máximo colapso articular. Luego, MediaPipe BlazePose extrajo coordenadas 3D de cadera, rodilla y tobillo para calcular el ángulo de desviación (FPPA) vía geometría vectorial para clasificar el riesgo de lesión.

Planteamiento del problema

En el deporte universitario de alto rendimiento, el Departamento de Fisioterapia de la Universidad Modelo, realiza evaluaciones biomecánicas durante las etapas de pretemporada, temporada y posttemporada. Dichas evaluaciones tienen como objetivo identificar factores de riesgo y prevenir lesiones graves en deportistas de básquetbol, principalmente, la ruptura del Ligamento Cruzado Anterior (LCA), cuya incidencia suele ser significativamente mayor en la población femenina (Bradshaw et al., 2025).

En entrevista con Juan Esteban Chan Narváez, encargado de la Clínica de Fisioterapia y Rehabilitación, se detectó un área de oportunidad, ya que actualmente se ha detectado una brecha en el proceso de evaluación, dado que, se presenta la imposibilidad de cuantificar con precisión el valgo dinámico de rodilla en el entorno clínico habitual. El experto del departamento, señala que, debido a que dicho fenómeno sucede en una fracción de segundo durante el aterrizaje y puede estar sujeto a cambios de dirección, la percepción del ojo humano resulta ineficiente para capturar el ángulo exacto de desviación o identificar el momento preciso del impacto máximo, también conocido como Peak Valgus (Chung et al., 2022).



Universidad Modelo

Escuela de Ingenierías – Ingeniería Biomédica

Protocolo de investigación

Título: Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla

Actualmente, si bien existen dispositivos de alta tecnología para medir dichos ángulos (Francia, 2024), el departamento de fisioterapia no cuenta con ellos. Esta ausencia de herramientas, obliga al personal de la clínica a depender de observaciones clínicas, o, en su defecto, a omitir este parámetro en los reportes de rendimiento. Sin un dato cuantitativo y objetivo, se complica el diseño de planes de entrenamiento preventivo personalizados y el seguimiento del rendimiento veraz del estado físico de las atletas.

Por lo tanto, existe la necesidad de desarrollar una herramienta automatizada basada en visión por computadora para la identificación y clasificación automática del valgo dinámico de rodilla en deportistas.

Justificación

La presente investigación surge como una propuesta para mejorar la evaluación clínica de las atletas de la Universidad Modelo y a evitar la incidencia desproporcionada de lesiones de LCA, el cual es un factor crítico que compromete la continuidad deportiva y el bienestar académico de las estudiantes (Bradshaw et al., 2025).

El desarrollo de este software responde a la necesidad de contar con herramientas de medición objetiva que puedan integrarse de manera ágil en el entorno clínico. Dado que el departamento no dispone actualmente de sistemas fijos, la adopción del marco de trabajo MediaPipe se fundamenta en su capacidad de ofrecer un equilibrio óptimo entre velocidad de procesamiento y exactitud en la estimación de esqueletos 3D utilizando únicamente visión artificial (Chung et al., 2022). Al prescindir de sensores físicos o marcadores, esta tecnología permite al Departamento de Fisioterapia obtener métricas cuantitativas críticas, como el Peak Valgus, sin interferir en la biomecánica natural de la atleta y buscando resultados reproducibles directamente en el lugar de la evaluación (Francia, 2024).

Estado del arte

El paso de sistemas de captura de movimiento basados en marcadores a soluciones "markerless" han representado un gran avance en la biomecánica para estudios de índole clínica o deportiva. Diversos estudios han demostrado que dichos sistemas ofrecen estimaciones cinemáticas similares a los métodos tradicionales. Un ejemplo de ello, es el estudio realizado por Song et al. (2023), en el cual se evaluaron ocho movimientos funcionales y se encontraron que las estimaciones markerless cinemáticas y cinéticas de las extremidades inferiores, presentaron resultados comparables a los sistemas con marcadores, principalmente en los puntos de rodilla y tobillo, mientras que hubo discrepancias en la estimación de la cadera.

Entre las herramientas utilizadas para análisis biomecánico en la visión por computadora, tenemos a MediaPipe, la cual, gracias a su capacidad de operar en tiempo real utilizando cámaras



convencionales ha ganado relevancia. Por lo mismo, en el estudio realizado por Francia, et al. (2024), se exploró el uso de esta herramienta para rehabilitación asistida por realidad virtual, destacando su potencial para el análisis de movimiento en ambientes no controlados. De igual manera, Akturk et al. (2026) propusieron un método de estimación de ángulos articulares basado en MediaPipe, el cual redujo significativamente los requerimientos de calibración tradicionales, facilitando además el cálculo de momentos articulares.

A pesar de sus ventajas operativas, la precisión de los sistemas markerless continúa dependiendo de distintos factores. Ruescas-Nicolau et al. (2024) analizaron qué variables afectan el error en aplicaciones biomecánicas basadas en deep learning, incluyendo la calidad de los keypoints, el tipo de movimiento y la arquitectura del modelo utilizado. Estos hallazgos coinciden con investigaciones previas como las de Stenum et al. (2021), quienes demostraron que la estimación de pose en video 2D puede utilizarse eficazmente para el análisis de la marcha humana.

Además de la estimación espacial del movimiento, en investigaciones recientes se han comenzado a integrar modelos de aprendizaje secuencial para interpretar la dinámica temporal de los gestos humanos con el uso de MediaPipe, tal como realizaron Varshini y Rukmani (2025), quienes propusieron una arquitectura híbrida basada en MediaPipe y redes Long Short-Term Memory (LSTM) para la detección de gestos en tiempo real. Esto para modelar dependencias temporales dentro de las secuencias de movimiento, con el fin de mejorar la interpretación continua de acciones dinámicas.

Solución Propuesta

La solución consiste en un ecosistema de software que automatiza la evaluación clínica del gesto motor Single Leg Drop Jump eliminando la subjetividad del observador humano. El sistema utiliza una arquitectura híbrida, en primera instancia, una Red Neuronal Convolucional (CNN ResNet50) acoplada a una red de memoria temporal (LSTM) procesa de forma automatizada la serie temporal del video para auto-sincronizar y aislar el fotograma exacto del colapso articular, eliminando la necesidad de una selección manual por parte del usuario. Una vez extraído este cuadro crítico, el núcleo del sistema aplica la arquitectura BlazePose de MediaPipe, la cual permite una captura de movimiento markerless (sin marcadores físicos) sobre la imagen aislada, facilitando su uso en cualquier entorno clínico sin necesidad de infraestructura costosa (Francia, 2024). Finalmente, el sistema transforma estas coordenadas espaciales en un valor angular cuantitativo y entrega un diagnóstico instantáneo basado en tres niveles de riesgo (Bajo, Moderado, Alto), permitiendo al preparador físico o fisioterapeuta tomar decisiones preventivas basadas en datos técnicos robustos y no en apreciaciones visuales.



Objetivos

2.1 Objetivo General

Diseñar e implementar un sistema basado en deep learning para la medición del ángulo valgo dinámico en imágenes de evaluaciones clínicas a deportistas de alto rendimiento de la Universidad Modelo.

2.2 Objetivos Específicos

- Estandarizar un protocolo de adquisición de video que optimice la detección de puntos anatómicos clave mediante el control de variables de entorno (distancia, altura y frecuencia de muestreo).
- Implementar una arquitectura de aprendizaje profundo en cascada (CNN-LSTM) sobre la secuencia temporal para el aislamiento automatizado del fotograma crítico de máximo colapso articular de forma objetiva (Francia, 2024).
- Implementar un algoritmo de inferencia de pose basado en MediaPipe para la extracción de coordenadas espaciales 3D (x, y, z) de los segmentos de cadera, rodilla y tobillo (Chung et al., 2022).
- Categorizar el nivel de riesgo de lesión mediante un motor de reglas basado en los rangos angulares clínicos validados por expertos.

Metodología

5.1 Captura de Datos y Protocolo de Video

Se registran videos de deportistas (20-26 años) realizando el salto monopodal (Single Leg Drop Jump). Para mitigar el error de distorsión geométrica por perspectiva y paralaje en el plano coronal, la cámara se posicionará de forma frontal a una distancia estandarizada de 2 metros y una altura fija de 0.9 metros, coincidiendo con el eje articular de la rodilla, siguiendo las recomendaciones de Francia (2024). La captura se configurará a una frecuencia de 60 fps para garantizar la resolución temporal necesaria. Como criterio de control experimental, los sujetos utilizarán pantalón corto ajustado para minimizar oclusiones visuales del tejido textil y reducir los falsos positivos en los algoritmos de detección geométrica posteriores.

5.2 Etiquetado Clínico y Creación del Dataset (Ground Truth)

Para el entrenamiento supervisado y la validación cruzada del sistema, se construirá un conjunto de datos estandarizado basado en el criterio de un experto. Un especialista en fisioterapia analizará



los videos de forma secuencial y etiquetará manualmente la marca de tiempo (timestamp) y el frame exacto donde ocurre el máximo colapso articular en valgo (Rango de Movimiento máximo o ROM). Esta tabla de datos estructurada conformará el Ground Truth del sistema, garantizando que el aprendizaje y la optimización del modelo computacional se fundamenten exclusivamente en evaluaciones clínicas profesionales y objetivas (Francia, 2024).

5.3 Preprocesamiento y Detección Temporal del Evento (Red Neuronal Primaria)

El sistema implementará una arquitectura de aprendizaje profundo enfocada en la detección de eventos temporales. Siguiendo las metodologías del estado del arte en normalización visual, cada fotograma se someterá primero a una fase de recorte y alineación espacial automática sobre los miembros inferiores para eliminar el ruido del fondo y estandarizar la escala del sujeto.

Posteriormente, las secuencias de imágenes procesadas alimentarán una red neuronal híbrida compuesta por una Red Neuronal Convolutiva (CNN) decapitada (ResNet50), esta red se propone debido a sus características de detección de siluetas, para la extracción de características geométricas espaciales, conectada en serie a una red de Memoria a Corto y Largo Plazo (LSTM) para el análisis secuencial. La CNN convolucionará cada frame reduciendo la dimensionalidad de los píxeles a un vector de características de alto nivel, mientras que la LSTM procesa la evolución temporal de estos vectores combinándolos con la tabla del Ground Truth. La salida (output) exclusiva de esta red primaria será el aislamiento automático del índice del frame exacto en el que sucede el valgo máximo, descartando el resto de la secuencia de video para optimizar la carga computacional subsiguiente.

5.4 Inferencia de Pose y Cálculo Biomecánico

Una vez extraída la imagen estática crítica por la red temporal, se aplicará MediaPipe Pose impulsado por el modelo de topología esquelética BlazePose como motor de inferencia espacial (Markerless Motion Capture). Al ejecutarse sobre un único fotograma aislado, el modelo localizará con alta fidelidad las proyecciones bidimensionales (x, y) de la cadera (cresta ilíaca), rodilla (centro rotuliano) y tobillo (maléolo externo).

La cuantificación del ángulo de proyección en el plano frontal (FPPA) se determinará mediante análisis geométrico vectorial. Se aplicará la fórmula del arco coseno sobre el producto escalar de los vectores anatómicos formados por el fémur (v_1) y la tibia (v_2):

$$\theta = \arccos\left(\frac{v_1 \cdot v_2}{\|v_1\| \|v_2\|}\right)$$

5.5 Clasificación de Riesgo



Universidad Modelo

Escuela de Ingenierías – Ingeniería Biomédica

Protocolo de investigación

Título: Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla

El ángulo de desviación final calculado a partir de la imagen aislada se restará del eje de alineación ideal para ser categorizado mediante un árbol de decisión determinista basado en los estándares clínicos de severidad para el riesgo de lesión del ligamento cruzado anterior (LCA):

Clasificación	Ángulo de desviación final
Bajo o Alineación normal	< 10 de desviación
Moderado o valgo leve	Entre 11 y 15 de desviación
Alto o valgo significativo	> 15 de desviación

Técnica de recolección de datos

La toma de videos para la creación del dataset se realizará en la clínica de fisioterapia de la Universidad Modelo. Para garantizar la uniformidad de las mediciones, se seguirá el siguiente protocolo de captura

- Equipo de captura: se grabará con un trípode posicionado a 2 m de distancia del sujeto y a una altura de 0.9 m.
- Parámetros del video: se utilizará una resolución mínima de 1080p a 60 FPS.
- Plano de captura: La cámara se situará en el plano frontal, asegurando que el eje óptico sea perpendicular a la zona de aterrizaje.
- Descripción del ejercicio: Se grabará la ejecución del Single Led Drop Jump, en el cual, la atleta se dejará caer desde un cajón de 30 cm, aterrizará e inmediatamente después del contacto con el suelo realizará un salto vertical lo más alto que le sea posible. En el primer contacto es donde el software buscará el ángulo del valgo.

Especificaciones para el dataset:

- Muestra: Estará conformada por atletas del sexo femenino en un rango de edades de 18 a 25 años de los equipos representativos de la universidad.
- Vestimenta: Las participantes deberán portar ropa deportiva ajustada, preferentemente shorts, de colores que contrasten con el fondo, permitiendo que las articulaciones de cadera, rodilla y tobillo sean visibles para facilitar la detección de los puntos claves.
- Iluminación: Se mantendrá una iluminación constante de tipo de luz blanca, para evitar sombras que puedan afectar la detección de puntos.
- Sonido ambiental: Con el fin de estandarizar las condiciones de ejecución y evitar distracciones que alteren la respuesta motora de las atletas, las pruebas se realizarán en un entorno controlado con ausencia de música o ruido ambiental excesivo.

Consideraciones éticas



Universidad Modelo

Escuela de Ingenierías – Ingeniería Biomédica

Protocolo de investigación

Título: Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla

Debido a que el sistema procesará registros de videos de atletas universitarias, se establecen los siguientes lineamientos para la protección de privacidad y datos de las atletas:

- Consentimiento informado: El uso de los videos cuenta con la autorización del encargado del Departamento de Fisioterapia de la Universidad Modelo, donde los atletas han otorgado previamente su consentimiento informado como pacientes de la clínica.
- Privacidad: Se garantizará el derecho a la privacidad de las atletas al mantener el anonimato de los registros. Ya que, el procesamiento a través de MediaPipe se enfocará exclusivamente en los puntos clave articulares, sin almacenar ni mostrar el rostro de las participantes.
- Manejo de datos: Los videos serán utilizados únicamente para el desarrollo del algoritmo y la validación del sistema, es decir, solo el equipo investigador y el personal clínico encargado tendrá acceso a los mismos, asegurando que la información no sea utilizada para fines ajenos a los objetivos de este protocolo.

Previsiones a futuro

Crear un sistema o experimento repetitivo para el usuario final, en este caso el departamento de fisioterapia y rehabilitación, para que pueda analizar el valgo de rodilla en los deportistas, de modo que, la calidad de los videos sea el indicado para que la red pueda analizar dicho video.

Propuesta de materiales:

- 2 cámaras de 60 fps y 1080p: esto garantiza la calidad y velocidad de imagen en todos los análisis y no tener diferencias en cuanto a la imagen.
- 2 trípode: esto es para la modulación de altura de las cámaras y sea siempre igual

Interacción con el usuario.

- Interfaz gráfica: crear una interfaz para que el usuario pueda analizar el valgo de rodilla subiendo los videos de los deportistas.

Avances tentativos

Los avances presentes tienen como objetivo principal presentar las metodologías o herramientas tentativas a utilizar, no representan el resultado final de dicho experimento, sino que sirven como indicio a las pruebas antes realizar las tomas con los sujetos de prueba finales (deportistas de la Universidad Modelo).

- **Creación de dataset**

Para la creación de dataset de prueba, se contó con la participación de 5 estudiantes universitarios. La toma de los videos se realizó a una distancia entre el sujeto de prueba y la cámara de 0.97m y



Universidad Modelo

Escuela de Ingenierías – Ingeniería Biomédica

Protocolo de investigación

Título: *Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla*

una altura de 0.80m, se hicieron 20 repeticiones por pierna, dando un total de 40 videos por persona, teniendo un total de 200 videos.

My Drive > Data_set_Valgo_de_Rod...

Type People Modified Source

Name	Owner	Date modified	File size
P5_Pierna_dch	me	10:27 PM	—
P5_Pierna_izq	me	10:27 PM	—
P4_Pierna_izq	me	8:19 PM	—
P4_Pierna_dch	me	8:18 PM	—
P3_Pierna_dch	me	May 6	—
P3_Pierna_izq	me	May 6	—
P2_Pierna_dch	me	May 6	—
P2_Pierna_izq	me	May 6	—

Imagen 1. Dataset.



Imágenes 2 y 3. Ejemplos de las tomas de video realizadas.

- Análisis de video mediante algoritmo BlazePose (MediaPipe)

Se procesó el video mediante un algoritmo predefinido, experimentando y validando los datos que nos brindaba el software, donde pudimos resaltar características importantes, entre ellas, el largo de extremidades. Cabe mencionar que, al tratarse de píxeles e intensidades, la veracidad de ese largo puede no ser siempre el óptimo debido a ligeros cambios provocados por movimientos involuntarios, lo que puede modificar la perspectiva del mismo.



Universidad Modelo

Escuela de Ingenierías – Ingeniería Biomédica

Protocolo de investigación

Título: Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla

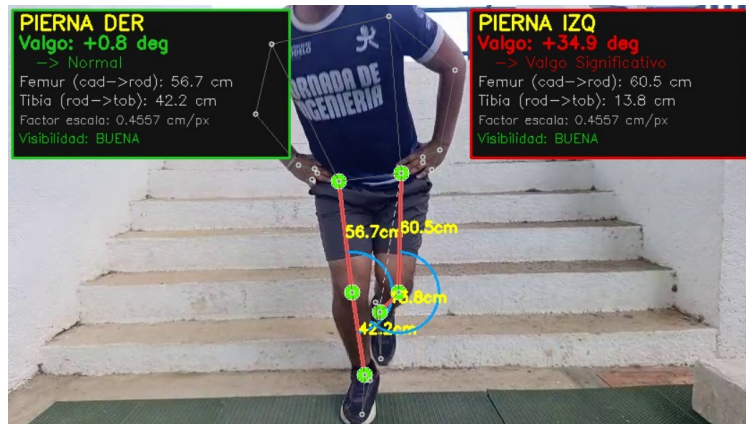


Imagen 4. Primeras pruebas con BlazePose (MediaPipe), donde se pueden visualizar la detección de los puntos de interés, la distancia de los vectores v1 y v2 por cada pierna y el ángulo del valgo.

Referencias

- Bradshaw, F. E., Gompels, B. D., Mainwaring, E., Tooth, G., Macmillan, A., Gupte, C., McDonnell, S. M., & Bailey, M. E. A. (2025). *Research priorities for ACL injuries in at-risk groups – The female population. The Knee*, 52. <https://doi.org/10.1016/j.knee.2025.02.001>
- Chung, J. L., et al. (2022). *Comparative analysis of skeleton-based human pose estimation. Future Internet*, 14(12).
- Francia, C. (2024). *Validation of a MediaPipe system for markerless motion analysis during virtual reality rehabilitation (Executive summary of the thesis)*. Politecnico di Milano.
- Song, K., Hullfish, T. J., Silva, R. S., Silbernagel, K. G., & Baxter, J. R. (2023). *Markerless motion capture estimates of lower extremity kinematics and kinetics are comparable to marker-based across 8 movements. Journal of Biomechanics*, 157, 111751. <https://doi.org/10.1016/j.jbiomech.2023.111751>
- Akturk, S., Derdiyok, F. B., & Serbest, K. (2026). *Markerless joint angle estimation using MediaPipe with a rapid setup for joint moment calculation. Multimedia Tools and Applications*, 85, 38. <https://doi.org/10.1007/s11042-026-21256-z>
- Ruescas-Nicolau, A. V., et al. (2024). *A deep learning model for markerless pose estimation based on keypoint augmentation: What factors influence errors in biomechanical applications? Sensors*, 24(6), 1923. <https://doi.org/10.3390/s24061923>
- Stenum, J., Rossi, C., & Roemmich, R. T. (2021). *Two-dimensional video-based analysis of human gait using pose estimation*.



Universidad Modelo

Escuela de Ingenierías – Ingeniería Biomédica

Protocolo de investigación

Título: Desarrollo de Software para la Detección Automatizada y Cuantificación Vectorial del Valgo Dinámico de Rodilla

- Varshini, T. S., & Rukmani, P. (2025). *MP-GestLSTM: Real time gesture detection using MediaPipe and LSTM*. *Systems Science & Control Engineering*.
- Malashin, I., et al. (2024). *Applications of Long Short-Term Memory (LSTM) networks in polymeric sciences: A review*. *Polymers*, 16(18).