

Implementación de un Sistema de Detección de Somnolencia en Conductores Mediante Visión por Computadora

Miguel Santiago Monsalve

Ángela Valeria Pinzón

Iván Ramiro Pinzón

Andrés Felipe Yañez

I. CONFORMACIÓN DEL EQUIPO

- **Andrés Felipe Yañez Villarraga** - Líder de proyecto: es responsable de supervisar la planificación, ejecución y finalización del sistema de detección de somnolencia en conductores. Entre sus responsabilidades principales se encuentran la coordinación del equipo, la gestión del cronograma del proyecto, y la asignación de tareas según las habilidades de cada miembro. Además, se encarga de revisar los hitos técnicos, asegurar que los entregables cumplan con los estándares de calidad, y mantener la comunicación con los stakeholders. No cuenta con experiencia previa respecto a proyectos de Computer Vision.
- **Miguel Santiago Monsalve Sánchez** - Investigador Principal: encargado de liderar el desarrollo técnico y científico del sistema de detección de somnolencia. Sus responsabilidades incluyen la investigación de técnicas avanzadas de visión por computadora, la selección y adaptación de algoritmos de detección, y la implementación de modelos de inteligencia artificial para el análisis de señales faciales y oculares. Miguel posee un doctorado en Ciencias de la Computación, con especialización en visión por computadora y aprendizaje automático. Su trayectoria incluye más de 10 publicaciones en revistas y conferencias internacionales, y ha trabajado en proyectos de investigación que abarcan desde la detección de emociones hasta el análisis del comportamiento humano en tiempo real.
- **Ángela Valeria Pinzón Cruz** - Desarrol-

ladora principal: responsable de transformar las ideas y teorías de investigación en soluciones funcionales y escalables. Entre sus responsabilidades clave se encuentran la programación y la implementación de los algoritmos de visión por computadora en el sistema final, la integración de hardware y software, y la optimización del sistema para su uso en tiempo real. Ángela es experta en arquitectura de software y desarrollo de sistemas complejos, con más de 6 años de experiencia en la creación de aplicaciones robustas y escalables en sectores que abarcan desde la seguridad hasta la automatización. Su conocimiento avanzado de frameworks como TensorFlow y OpenCV, así como su habilidad en lenguajes de programación como Python y C++, le permite liderar la fase de desarrollo del proyecto con una combinación de eficiencia y precisión

- **Iván Ramiro Pinzón Pinto** - experto en datos: encargado de gestionar, procesar y analizar los grandes volúmenes de datos generados durante el desarrollo del sistema de detección de somnolencia. Su rol incluye la limpieza y preprocesamiento de datos, la selección y ajuste de características relevantes para los modelos de aprendizaje automático, y la evaluación del rendimiento de los algoritmos en función de los datos recogidos. Iván cuenta con más de 5 años de experiencia en ciencia de datos, especializándose en la manipulación y análisis de grandes conjuntos de datos en tiempo real. Es experto en el uso de herramientas como Python, Pandas,

NumPy, y TensorFlow, y tiene un profundo conocimiento de técnicas avanzadas de modelado predictivo y aprendizaje supervisado. Su capacidad para extraer información significativa de los datos y su conocimiento de las mejores prácticas en el manejo de datos aseguran que el sistema esté respaldado por un análisis robusto y bien fundamentado.

II. MARCO TEÓRICO

A. Descripción del Problema

La somnolencia al volante es una de las principales causas de accidentes de tráfico en todo el mundo, contribuyendo a una tasa de mortalidad que varía entre el 3% y el 30%, además de causar un alto número de lesiones graves en las carreteras (BaHammam et al., 2023). Detectar y prevenir episodios de somnolencia en los conductores es un desafío crucial para mejorar la seguridad vial (Ahmed et al., 2023).

El **objetivo principal** de este proyecto es **replicar y evaluar un sistema de detección de somnolencia basado en el seguimiento de ojos y boca utilizando técnicas de visión por computadora**. Si bien el seguimiento ocular y bucal es un enfoque bien establecido, este proyecto busca identificar y abordar las deficiencias que puedan surgir al aplicar este método en condiciones reales de conducción.

Además, aspiramos a expandir este enfoque básico integrando otros indicadores fisiológicos y comportamentales, como el ritmo cardíaco y los movimientos de la cabeza, siempre que el tiempo y los recursos disponibles lo permitan. Sin embargo, es importante destacar que el objetivo inicial es la replicación y optimización del modelo basado en el seguimiento ocular y bucal, mientras que las implementaciones adicionales serán exploradas como posibles mejoras durante las fases posteriores del proyecto.

Este enfoque progresivo nos permitirá establecer una base sólida sobre la cual construir un sistema más robusto, capaz de operar en tiempo real y bajo una variedad de condiciones de conducción. Entre los desafíos técnicos que prevemos se encuentran la gestión de diferentes condiciones ambientales,

como variaciones de iluminación, y la adaptabilidad a las características individuales de cada conductor.

Limitaciones: Las principales limitaciones incluyen:

- **Variabilidad Individual:** Las señales de somnolencia pueden variar significativamente entre diferentes personas, lo que dificulta la creación de un modelo universalmente aplicable.
- **Condiciones Ambientales:** Las condiciones de iluminación y la posición del conductor pueden afectar la precisión del sistema. Por ejemplo, la baja iluminación o los reflejos en los lentes pueden interferir con el seguimiento ocular.
- **Rendimiento en Tiempo Real:** Dado que el sistema debe funcionar en tiempo real para ser efectivo, es necesario equilibrar la precisión con la velocidad de procesamiento.

B. Justificación del Problema

La detección de somnolencia en conductores es de gran relevancia debido a la cantidad de accidentes que se pueden prevenir con la intervención oportuna (Majeed et al., 2023). A nivel académico, este problema ofrece la oportunidad de avanzar en el campo de la visión por computadora aplicada a la seguridad vehicular, un área que continúa evolucionando rápidamente con la integración de inteligencia artificial y técnicas de Deep Learning (Pavel et al., 2022). Desde un punto de vista práctico, la solución tiene aplicaciones directas en la industria automotriz, especialmente en vehículos de conducción asistida y autónoma, donde la capacidad de monitorizar el estado del conductor puede llegar a ser vital. (Garikapati & Shetiya, 2024).

C. Impacto Potencial

Resolver este problema tiene el potencial de salvar vidas y reducir significativamente los costos asociados con accidentes de tráfico relacionados con la fatiga del conductor (BaHammam et al., 2023). A nivel industrial, una solución efectiva podría ser integrada en sistemas de seguridad avanzada de vehículos modernos, contribuyendo a la adopción de tecnologías inteligentes en la conducción (Jagatheesaperumal et al., 2024). En

la academia, esto abriría nuevas líneas de investigación en la intersección de la visión por computadora, la inteligencia artificial y la seguridad vial, especialmente utilizando técnicas de Deep Learning (Safarov et al., 2023). A nivel social, la implementación de este tipo de sistemas podría mejorar la conciencia sobre los peligros de la somnolencia al volante y promover hábitos de conducción más seguros (Yang & Yi, 2024).

III. RESUMENES Y ANÁLISIS

1. *Biosignals Monitoring for Driver Drowsiness Detection Using Deep Neural Networks*

Resumen: Este estudio investiga cómo las señales fisiológicas pueden utilizarse para detectar somnolencia en conductores mediante redes neuronales profundas. Se analizaron tres señales clave: HRV, EDA y el seguimiento ocular, utilizando modelos específicos para cada una.

- **HRV:** Emplearon un modelo secuencial (SNN) para interpretar patrones de frecuencia cardíaca, alcanzando una precisión del 98.28%.
- **EDA:** Usaron una red convolucional unidimensional (1D-CNN) para mapear cambios en la conductancia de la piel, alcanzando una precisión del 96.32%.
- **Seguimiento ocular:** Se utilizó un CRNN para analizar PERCLOS y parpadeos, pero tuvo problemas debido a un conjunto de datos desbalanceado.

Análisis Crítico: La estrategia de emplear modelos separados para cada señal fisiológica permite maximizar la precisión para cada caso específico. Sin embargo, esto puede limitar la capacidad del sistema para detectar correlaciones entre las señales, lo que podría aumentar su precisión. Los hallazgos sobre HRV y EDA son muy alentadores, pero los problemas con el seguimiento ocular resaltan la importancia de un mejor preprocesamiento de los datos, lo que mejoraría la efectividad del sistema en escenarios reales.

2. *Real-Time Deep Learning-Based Drowsiness Detection: Leveraging Computer-Vision and Eye-Blink Analyses for Enhanced Road Safety*

Resumen: Este artículo trata sobre la implementación de un sistema basado en visión por

computadora que analiza patrones faciales para detectar somnolencia en conductores en tiempo real. Se enfoca en la detección de parpadeos, ojos cerrados y abiertos, así como bostezos.

- **Resultados:** El sistema logró una precisión del 95.8% en la detección de ojos cerrados y del 97% en la detección de ojos abiertos, además de identificar caídas de cabeza relacionadas con fatiga severa.

Análisis Crítico: El uso de visión por computadora es eficaz para la detección en tiempo real y proporciona buenos resultados en entornos controlados. Sin embargo, el rendimiento puede verse afectado en condiciones de iluminación variables y diferencias faciales entre los usuarios. Integrar esta metodología con otras señales fisiológicas, como HRV o EDA, podría mejorar su robustez en escenarios más desafiantes.

3. *Drowsiness Detection Using Brain Signal Recognition Deep Neural Network (BSRDNN)*

Resumen: Este artículo examina el uso de redes neuronales profundas para analizar señales EEG y detectar somnolencia en los conductores. Se utilizó una combinación de redes convolucionales unidimensionales (1D-CNN) y redes LSTM para capturar la naturaleza secuencial de las señales cerebrales.

- **Resultados:** Se alcanzó una precisión del 94.8%, mostrando una alta capacidad para clasificar estados de alerta y somnolencia con precisión.

Análisis Crítico: La combinación de CNN y LSTM es apropiada para analizar la naturaleza secuencial de las señales EEG. Sin embargo, la naturaleza invasiva y costosa del EEG puede limitar su implementación en aplicaciones prácticas. Sería interesante investigar métodos menos invasivos o combinar EEG con otros métodos de detección para hacer más viable su uso en el mundo real.

4. *A Deep Learning Approach for Real-Time Detection of Sleep Spindles*

Resumen: Este documento aborda la detección de husos del sueño utilizando EEG y el modelo de red neuronal profunda SpindleNet. El sistema está diseñado para detectar husos en tiempo real, lo que

tiene implicaciones para la mejora del rendimiento cognitivo.

- **Resultados:** SpindleNet demostró alta precisión en la detección de husos del sueño, con una latencia de detección promedio de 150 ms, adecuada para aplicaciones en tiempo real.

Análisis Crítico: SpindleNet muestra ser una implementación eficaz para la detección de husos del sueño, particularmente al manejar datos ruidosos. No obstante, al igual que otros sistemas EEG, su viabilidad fuera del entorno de laboratorio es limitada por la necesidad de equipo especializado y su naturaleza invasiva. Sería beneficioso explorar técnicas más accesibles o sistemas híbridos que combinen EEG con otros indicadores.

5. Automated Detection of Sleep Stages Using Deep Learning Techniques: A Systematic Review of the Last Decade (2010–2020)

Resumen: Este artículo realiza una revisión sistemática de los avances en la detección de fatiga y somnolencia mediante técnicas de aprendizaje profundo y sensores fisiológicos. Se enfoca en la integración de HRV, EDA y señales visuales para una detección más precisa de la fatiga en conductores.

- **Metodología:** Utiliza una combinación de modelos de redes neuronales para analizar diferentes tipos de señales y ofrecer una evaluación más completa del estado del conductor.

IV. FUENTES DE DATOS

Para nuestro proyecto de detección de somnolencia en conductores, utilizaremos un conjunto de bases de datos públicas que han sido ampliamente empleadas en investigaciones similares. Estas bases de datos cubren una variedad de escenarios de somnolencia, características fisiológicas y condiciones de conducción en tiempo real.

Closed Eyes In The Wild (CEW)

- **Origen:** Closed Eyes In The Wild (CEW) fue diseñada para la detección de ojos cerrados en imágenes capturadas en situaciones no controladas.

- **Tamaño:** Este conjunto de datos contiene aproximadamente 24,000 imágenes etiquetadas de ojos cerrados y abiertos.
- **Características:** Las imágenes varían en condiciones de iluminación, ángulos y características faciales, lo que es ideal para entrenar modelos robustos de detección de somnolencia basada en parpadeos y ojos cerrados.

YawDD Dataset

- **Origen:** YawDD Dataset contiene dos conjuntos de datos de videos de conductores, diseñados principalmente para la detección de bostezos, así como para el reconocimiento y seguimiento de rostros y bocas.
- **Propósito:** El conjunto de datos está diseñado para probar algoritmos y modelos de detección de bostezos en condiciones reales y variadas de iluminación.
- **Tamaño:**
 - **Primer conjunto de datos:** Contiene 322 videos de conductores (hombres y mujeres) con características faciales variadas, incluidos conductores con y sin gafas/sombrillas, de diferentes etnias. Los videos están clasificados en tres situaciones: 1) conducción normal (sin hablar), 2) conducción mientras hablan o cantan, y 3) conducción mientras bostezan.
 - **Segundo conjunto de datos:** Contiene 29 videos en los que la cámara está instalada en el tablero del conductor. Cada video incluye escenas de conducción normal, conducción mientras hablan y conducción mientras bostezan.

UTA Real-Life Drowsiness Dataset (UTA-RLDD)

- **Origen:** UTA Real-Life Drowsiness Dataset (UTA-RLDD) fue creado por la Universidad de Texas en Arlington para la tarea de detección de somnolencia en múltiples etapas, abordando tanto casos extremos y visibles como casos sutiles donde las microexpresiones juegan un papel clave.
- **Propósito:** Este conjunto de datos tiene como objetivo detectar casos sutiles de somnolencia, que son importantes para activar mecanismos preventivos antes de que la somnolencia

se vuelva crítica. Las microexpresiones sutiles tienen fuentes fisiológicas e instintivas, lo que dificulta su simulación realista por actores.

- **Tamaño:** El conjunto de datos consiste en aproximadamente 30 horas de videos RGB de 60 participantes saludables. Cada participante grabó tres videos, uno para cada una de las siguientes clases: alerta, baja vigilancia y somnolencia, resultando en un total de 180 videos.
- **Características:**
 - Los participantes fueron estudiantes y personal universitario, de diferentes etnias y edades (de 20 a 59 años, con una media de 25 y una desviación estándar de 6). Los sujetos utilizaron gafas en 21 de los 180 videos y tenían barba en 72 de los videos.
 - Los videos fueron grabados en diversos entornos de la vida real, desde diferentes ángulos y fondos, utilizando cámaras de teléfonos móviles o cámaras web con una tasa de fotogramas menor a 30 fps, representando condiciones típicas de grabación.

NITYMED Dataset

- **Origen:** NITYMED Dataset fue capturado en Patras, Grecia, y contiene videos de conductores en condiciones de conducción nocturna, donde la detección de somnolencia es crucial.
- **Propósito:** Este conjunto de datos está diseñado para probar y comparar algoritmos y modelos de detección de somnolencia en condiciones nocturnas. También se puede utilizar para aplicaciones de seguimiento de rostro, boca y ojos.
- **Tamaño:** El conjunto de datos contiene un total de 130 videos divididos en dos categorías:
 - **Bostezo:** Los conductores bostezan tres veces en cada video, con una duración aproximada de 15-25 segundos (107 videos).
 - **Microsueños:** Los conductores hablan, miran a su alrededor y tienen microsueños en videos de aproximadamente 2 minutos de duración (21 videos).
- **Características:** El conjunto de datos incluye

videos de 21 participantes (11 hombres y 10 mujeres) con diversas características físicas (color de cabello, barba, gafas, etc.). Los videos fueron grabados en condiciones de iluminación nocturna natural con una leve mejora proporcionada por las luces interiores más bajas del vehículo, simulando condiciones de conducción en una avenida oscura y poco concurrida.

The Driver Drowsiness Dataset (DDD)

- **Origen:** The Driver Drowsiness Dataset (DDD) combina grabaciones de video y señales fisiológicas de conductores en entornos controlados y reales.
- **Tamaño:** Contiene datos de aproximadamente 60 participantes en diversas condiciones de conducción.
- **Características:** Los datos incluyen secuencias de video que muestran expresiones faciales (parpadeos, bostezos) y señales fisiológicas como EEG y ECG.

Descripción de los Datos

- **Número de Muestras:** Trabajaremos con un conjunto diverso de miles de imágenes y múltiples horas de video provenientes de las bases de datos mencionadas.
- **Características:** Los datos incluyen imágenes faciales, secuencias de video de ojos y boca, así como señales EEG y ECG en algunos casos. Estas características permitirán el análisis detallado de los patrones de somnolencia, como parpadeos y señales fisiológicas relacionadas.

Recopilación de Nuevos Datos

Actualmente, no planeamos recopilar nuevos datos para este proyecto, ya que los conjuntos de datos seleccionados proporcionan una cobertura amplia y adecuada para nuestras necesidades. Sin embargo, en el transcurso del proyecto, si se identifica una **necesidad específica** que no esté cubierta por los datos existentes, podríamos considerar la recopilación de nuevos datos, utilizando herramientas como cámaras web para capturar videos en tiempo real bajo diferentes condiciones de iluminación y posiciones del conductor.

Herramientas de Análisis y Preprocesamiento

Utilizaremos bibliotecas de Python como OpenCV para la interpretación de caras y extracción de características faciales, como la distancia entre ojos, labios y la postura de los hombros. Esto permitirá ajustar métricas y mejorar la precisión de la detección de somnolencia. El preprocesamiento incluirá la normalización de imágenes y el manejo de desbalanceo en las clases, lo que garantizará la calidad y consistencia de los datos utilizados en el entrenamiento del modelo.

V. METODOLOGÍA

1. Selección de Algoritmos

Redes Neuronales Convolucionales (CNNs): Para este proyecto, proponemos inicialmente utilizar MTCNN o Multi-Task Cascaded Convolutional Neural Networks es una red neuronal que detecta rostros y puntos de referencia faciales en imágenes (Zhang et al., 2016). MTCNN se utiliza principalmente para la detección de rostros y la alineación facial. Es capaz de detectar rostros en imágenes, junto con puntos clave (ojos, nariz, boca), lo que facilita la identificación de la pose y la alineación de las caras.

Redes Neuronales Recurrentes (RNNs) con LSTM: Para manejar las secuencias de video y los cambios temporales en las señales fisiológicas, utilizaremos una arquitectura de LSTM (Long Short-Term Memory). Las LSTM se utilizarán para analizar la secuencia de parpadeos, movimientos de cabeza, y otros cambios temporales en el comportamiento del conductor que son indicativos de somnolencia.

Depthwise Separable Convolution: Adicionalmente, proponemos utilizar MobileNet debido a su arquitectura ligera y eficiente, diseñada específicamente para dispositivos con recursos limitados como teléfonos móviles o sistemas embebidos (Howard et al., 2017). MobileNet utiliza convoluciones separables en profundidad, lo que reduce significativamente la cantidad de parámetros y operaciones, permitiendo un procesamiento rápido en tiempo real sin sacrificar demasiado la precisión, lo que la hace ideal para tareas como la detección de somnolencia en conductores donde la baja latencia es importante.

MobileNet es un modelo de visión por ordenador de código abierto de Google diseñado para entrenar clasificadores. Utiliza convoluciones en profundidad para reducir significativamente el número de parámetros en comparación con otras redes, lo que resulta en una red neuronal profunda ligera. MobileNet es el primer modelo de visión por ordenador móvil de Tensorflow.

2. Implementaciones Existentes

OpenCV: OpenCV se utilizará para detectar y rastrear características faciales clave, como los ojos, la boca y la posición de la cabeza. Estos datos servirán como entradas adicionales para las redes CNN y LSTM.

Integración de Múltiples Modalidades: Como se planteó en II - Marco Teórico, en las etapas avanzadas del proyecto, luego de conseguir nuestro **objetivo principal** la idea sería poder integrar datos de diferentes modalidades (imágenes, videos, señales fisiológicas) para crear un modelo multi-modal que pueda capturar señales de somnolencia de manera más holística.

VI. EVALUACIÓN DE RESULTADOS

1. Evaluación Cualitativa

La evaluación cualitativa se centrará en analizar cómo el modelo se comporta en situaciones del mundo real. Dado que somos un equipo de cuatro personas, nuestras pruebas estarán limitadas tanto en alcance como en diversidad de condiciones. Realizaremos pruebas en un conjunto limitado de situaciones controladas y no controladas, considerando variaciones como diferentes condiciones de iluminación y posiciones del conductor. Aunque estas pruebas serán representativas, reconocemos que estarán limitadas en comparación con estudios más amplios. Además, se considerará la posibilidad de obtener retroalimentación de usuarios y expertos para validar la efectividad del sistema en un contexto práctico.

2. Evaluación Cuantitativa

La evaluación cuantitativa utilizará las siguientes métricas estándar en machine learning para medir el rendimiento del modelo:

- **Precisión (Accuracy):** La fracción de predicciones correctas sobre el total de predicciones realizadas.

- **Precisión (Precision):** La fracción de verdaderos positivos sobre el total de predicciones positivas.
- **Recall (Sensibilidad):** La fracción de verdaderos positivos sobre el total de verdaderos positivos y falsos negativos.
- **F1-Score:** La media armónica de precisión y recall.
- **ROC-AUC (Área Bajo la Curva ROC):** Esta métrica evalúa la capacidad del modelo para distinguir entre clases (somnolencia vs alerta) en diferentes umbrales de decisión.

Analizaremos la matriz de confusión para comprender mejor dónde el modelo está cometiendo errores. También utilizaremos la validación cruzada k-fold para asegurar que el modelo es robusto y no está sobreajustado a un subconjunto específico de los datos.

3. Análisis Comparativo

Compararemos los resultados de nuestro modelo con otros modelos open source disponibles, así como con CNNs básicas y otros modelos simplificados. Este análisis comparativo nos permitirá evaluar si la complejidad adicional de nuestro modelo justifica las mejoras en el rendimiento. Es posible que implementemos técnicas de validación cruzada como k-fold cross-validation para obtener una estimación robusta del rendimiento del modelo en diferentes subconjuntos de datos.

REFERENCES

- Ahmed, M. I. B., Alabdulkarem, H., Alomair, F., Aldossary, D., Alahmari, M., Alhumaidan, M., Alrassan, S., Rahman, A., Youldash, M., & Zaman, G. (2023). A deep-learning approach to driver drowsiness detection. *Safety*, 9(3). <https://doi.org/10.3390/safety9030065>
- BaHammam, A. S., Hunasikatti, M., & Pandi-Perumal, S. R. (2023). Traffic safety in sleep deprivation, sleepiness, and sleep disorders. In A. S. BaHammam & M. Hunasikatti (Eds.), *Sleep apnea frontiers: Pathophysiology, diagnosis, and treatment strategies* (pp. 161–179). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-7901-1_11
- Garikapati, D., & Shetiya, S. S. (2024). Autonomous vehicles: Evolution of artificial intelligence and the current industry landscape. *Big Data and Cognitive Computing*, 8(4). <https://doi.org/10.3390/bdcc8040042>
- Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M., & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.
- Jagatheesaperumal, S. K., Bibri, S. E., Huang, J., Rajapandian, J., & Parthiban, B. (2024). Artificial intelligence of things for smart cities: Advanced solutions for enhancing transportation safety. *Computational Urban Science*, 4(1), 10. <https://doi.org/10.1007/s43762-024-00120-6>
- Majeed, F., Shafique, U., Safran, M., Alfarhood, S., & Ashraf, I. (2023). Detection of drowsiness among drivers using novel deep convolutional neural network model. *Sensors*, 23(21). <https://doi.org/10.3390/s23218741>
- Pavel, M. I., Tan, S. Y., & Abdullah, A. (2022). Vision-based autonomous vehicle systems based on deep learning: A systematic literature review. *Applied Sciences*, 12(14). <https://doi.org/10.3390/app12146831>
- Safarov, F., Akhmedov, F., Abdusalomov, A. B., Nasimov, R., & Cho, Y. I. (2023). Real-time deep learning-based drowsiness detection: Leveraging computer-vision and eye-blink analyses for enhanced road safety. *Sensors*, 23(14). <https://doi.org/10.3390/s23146459>
- Yang, E., & Yi, O. (2024). Enhancing road safety: Deep learning-based intelligent driver drowsiness detection for advanced driver-assistance systems. *Electronics*, 13(4). <https://doi.org/10.3390/electronics13040708>
- Zhang, K., Zhang, Z., Li, Z., & Qiao, Y. (2016). Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, 23(10), 1499–1503.