

# **Tesis de Maestría**

## **Reconocimiento de acciones humanas en video, utilizando detección de contornos y Deep Learning**

Cronograma de Trabajo

**Alumno: Rodrigo Fondato**

**Tutor: Juliana Gambini**

### **1. Descripción del problema**

Las técnicas para la detección automática de acciones realizadas por personas en capturas de videos pueden dividirse, a grandes rasgos, en aquellas que utilizan herramientas de procesamiento de imágenes para obtener descriptores espacio-temporales y el posterior uso de modelos clásicos de machine learning para la clasificación, y en aquellas basadas en la explotación de grandes volúmenes de datos y la consecuente utilización de modelos de deep learning para etiquetar las acciones detectadas.

Éstos últimos modelos, más investigados en la actualidad, permiten obtener descriptores más robustos que los primeros, pero requieren una gran cantidad de recursos computacionales (poder de procesamiento y memoria), tiempo y disponibilidad de datos para su entrenamiento, debido al tamaño de los archivos de video en memoria y la cantidad de parámetros que deben ser optimizados. Existe además, un compromiso entre el horizonte temporal y la resolución espacial. Es decir, a mayor resolución de la imagen de entrada, menor cantidad de cuadros podrán ser analizados por el modelo para efectuar la clasificación, y viceversa.

El problema que se intenta resolver es, para una determinada resolución de video de entrada, la expansión del horizonte temporal y la disminución de la

cantidad de parámetros de un modelo basado en deep learning, con la consecuente reducción de la cantidad de datos y recursos computacionales necesarios para el entrenamiento, obteniendo a la vez las ventajas de robustez y rendimiento que otorgan éste tipo de modelos, específicamente para el ámbito de reconocimiento de acciones humanas básicas.

## **2. Objetivos específicos**

Se propone realizar las siguientes actividades:

1. Definición del proceso de gestión del proyecto, selección de herramientas a utilizar para la gestión y creación de lista de tareas.
2. Definición del proceso de revisión de la literatura y fuentes a investigar, armado de consultas y criterios de selección de artículos, y ejecución del proceso de selección. Lectura de los artículos, y armado del estado de la cuestión.
3. Identificación, clasificación, selección y obtención de bases de datos públicas que contengan segmentos de video que muestren acciones humanas.
4. Análisis y selección de los segmentos de video según categorías, calidad, duración y formato.
5. Implementación y ejecución de un proceso para adaptar y homogeneizar los distintos archivos, obteniendo como resultado una única base de datos.
6. Diseño e implementación del modelo de reconocimiento de acciones humanas.
7. Diseño de las métricas e implementación de la herramienta que las obtendrá.
8. Entrenamiento y optimización del modelo.
9. Revisión del diseño de las características y el modelo, propuesta e implementación de mejoras. Obtención de nuevas métricas y elección del modelo ganador.
10. Instalación y ejecución de modelos pre-entrenados existentes. Obtención de métricas.
11. Análisis de los resultados obtenidos en la comparación de modelos. Elaboración de conclusiones y finalización del documento de tesis.

[illegible]