**Infant Drowning Prevention and Alert System using Video Vision Transformer**

**Abstract**

Drowning deaths represent the third leading cause of accidental deaths worldwide. Drowning accidents in indoor swimming pools are growing in recent years, especially for children. This is because traditional techniques for the supervision and care of people, especially children, in large pools are inefficient or, in some cases, non-existent. Nowadays, this problem has become a topic of interest for several researchers who seek to propose different methods of drowning detection. This project seeks to propose the process to be followed to develop a drowning detection system in swimming pools using Video Vision Transformer. This project proposes a novel embedding scheme and a number of Transformer variants to model video clips. This model extracts spatio-temporal tokens from the input video, which are then encoded by a series of transformer layers. The output tokens of the pre-trained ViViT contain spatio-temporal information of drowning scenes. Then transform a query scene and candidate scenes into output token features using the pre-trained ViViT and calculate the similarity between the tokens with cosine similarity. The proposed network is designed to be lightweight based on the Temporal Transformer and Feature Pyramid Networks to detect drowning infants underwater instead of large networks. Therefore, the creation of automated systems that are able to provide with accurate alarms when certain events take place is of paramount importance, as this can heavily reduce swimming pool drowning accidents and improve the efficiency of the system. The experimental results show that the proposed network can ensure relatively good accuracy with fast detection speed.

**Technology Stack**

* Server Side : Python 3.7.4(64-bit) or (32-bit)
* Client Side : HTML, CSS, Bootstrap
* IDE : IDLE, Flask 1.1.1
* Back end : MySQL 5.
* Server : Wampserver 2i
* DL DLL : TensorFlow, Pandas, SiKit Learn