**Proposal of changes for thesis.**

**Thesis topic: A safety camera solution to alert risk for children at home.**

**Motivation:**

* Detecting risk sources at home for children is essential to ensure their safety and well-being, as homes often harbor hidden hazards that can lead to accidents or injuries. Everyday items like sharp objects, uncovered electrical outlets, toxic cleaning agents, or unstable furniture can pose significant risks to curious and active children.
* A common approach is to use safety cameras to keep track at the children.
* However, conventional cameras is only valueable when there is someone who always monitor it, which is unrealistic.
* The issue can be handled by applying automation in tracking and alert sending by using Deep learning models to for Detection and Ranging.
* This is the reason the thesis topic is proposed.

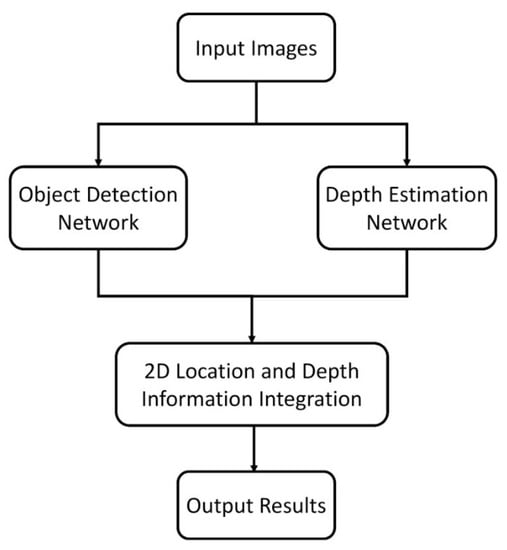
**Problem description:**

* In the alert system, the most important part is to detect whether the children is near the risk sources e.g. smoke, fire, hot stove, knives, or basically in a location without supervision for a period of time.
* The main components of the system shall be:
  + The data as images from security cameras
  + The detection component to detect the children and risk sources
  + The ranging component to detect distance between the children and risk source.
  + The evaluation and alert component to inform parents in case risks detected.
* In actual scenarios, there might be cases where further risks need to be analyzed as stranger approach, window not closed, children’s unsafe poses,… **but in scope of this thesis, the author will focus on analyzing distance between the children and risk source only**. This limited scope ensure available non-privacy data and still meaningful in evaluation of the problem.

**Relevant research and analysis:**

**Note**: The information is not yet formatted, will be formatted later

* “DOES DEPTH ESTIMATION HELP OBJECT DETECTION?” - Bedrettin Cetinkaya.
  + Description: estimated depth, effects of different state-of-the-art depth estimation networks, effects of using different indoor and outdoor RGB-D datasets as training data for depth estimation, and different architectural choices for integrating depth to the base object detector network.
  + A diagram of a computer system

    Description automatically generated
  + Analysis: mainly focus on using estimated depth map to predict the object label.
* “Object Detection and Depth Estimation Approach Based on Deep Convolutional Neural Networks” –
  + Description: object detection and depth estimation approach based on deep convolutional neural networks (CNNs). Improve object detection through the incorporation of transfer connection blocks (TCBs), in particular, to detect small objects in real time. For depth estimation, we introduce binocular vision to the monocular-based disparity estimation network, and the epipolar constraint is used to improve prediction accuracy.
  + 
  + Analysis: Depth map data is always calculated, which is not really efficiency for computation resource. Also, training dataset requires stereo vision data which is not always available for private data such as home safety.

**Objectives of the thesis:**

* Gain comprehensive understanding of detection and depth estimation models and their application home safety.
* Implement solution based on camera, using depth estimation models and object detection models to evaluate risks for children at home and alert the parents.
* Identify suitable benchmarks for evaluation.

**Scopes of the thesis:**

* The thesis primarily concentrates on techniques that enable fast object detection and depth estimation in in-door scenarios.
* The datasets:
  + For object detection: re-label the available dataset from ImageNet, to further detect children, classify risk sources.
  + For depth estimation: available popular indoor datasets NYU, DIODE shall be used.
* The goal is to develop an efficient real-time application to determine whether the children is at risk and report to parents.
  + This shall be demonstrated by several youtube videos on risk cases on children at home.

**Draft solutions:**

* Both solutions shall be used and tested to compare accuracy and effectiveness:

Solution 1:

* + Use 2 separate models:
    - YOLO model to detect the risk source and chidren.
    - Only if children is detected in stage 1, a monocular depth estimation model shall be run to find distance between the children and the risk source

Solution 2:

* + Leverage yolo-mde with different head-backbone to detect both objects and depth at same time.
  + Difficulties:
    - Lack of data, especially data with both depth and object detection
* The evaluation benchmark shall be separated for object detection and depth estimation.
* The main contributions shall be:
  + Re-labeled dataset for specific use case
  + Comparision between compatibility of solutions in YOLO-MDE (backbone/head) for this specific use case. (Error, run time)