PROJECT PHASE 2
GROUP 12

- **What problem will your Computer Vision solution solve, and for whom?**

  In a multi view or multi scale scenario, we tend to detect multiple instances of the same object. At times, this leads to double counting or redundancy for downstream processing modules. One example can be detecting damages in vehicles from images captured from different angles.

- **What value will it provide them? What are their pain points?**

  The repair cost can be exaggerated if the solution is not able to recognize that the two different images point to the same damage. The main pain point will be to identify duplicate instances of the same object, when captured from multiple angles. Our approach could also be used as a foundation for many other applications.

- **How big is the potential market?**

  Identifying duplicate instances of the same Object could be used in most of the vehicle insurance coverage companies. Which is a big market all over the world, right now. Since there are many similar problems that can be addressed with our approach, the market could be even bigger.

- **What open source code is available that is relevant to your topic?**

  While searching through papers that are relevant to our project, found following codes that could be useful.
  Finding the exact rotation between two images independently of the translation:
  Found following codes relevant to the paper.
  https://github.com/laurentkneip/opengv/blob/master/doc/addons/mainpage.dox

  Deep Fundamental Matrix Estimation:
  The below is the work done by the people at INTEL Labs, which is published in the paper "Deep Fundamental Matrix Estimation".
  https://github.com/intel-isl/DFE

  RANSAC Algorithm:
  The below is the code for RANSAC Algorithm, which could be used to match the in-liners across images.
  https://github.com/marktao99/python/blob/master/CVP/samples/ransac.py

- **What data is available for testing and/or training algorithms?**

  Industrial partner (GenPact) has provided the Test Data relevant to this project. The data set contains 7,200 images of 100 objects. Each object was turned on a turntable through 360 degrees to vary object pose with respect to a fixed color camera. Images of the objects were taken at pose intervals of 5 degrees. This corresponds to 72 poses per object. Their images were then size normalized. Objects have a wide variety of complex geometric and reflectance characteristics.

Original data source and banner. image:
http://www1.cs.columbia.edu/CAVE/software/softlib/coil-100.php

- **Is labeled data available? How much? How is the data licensed? Is it under copyright protection?**

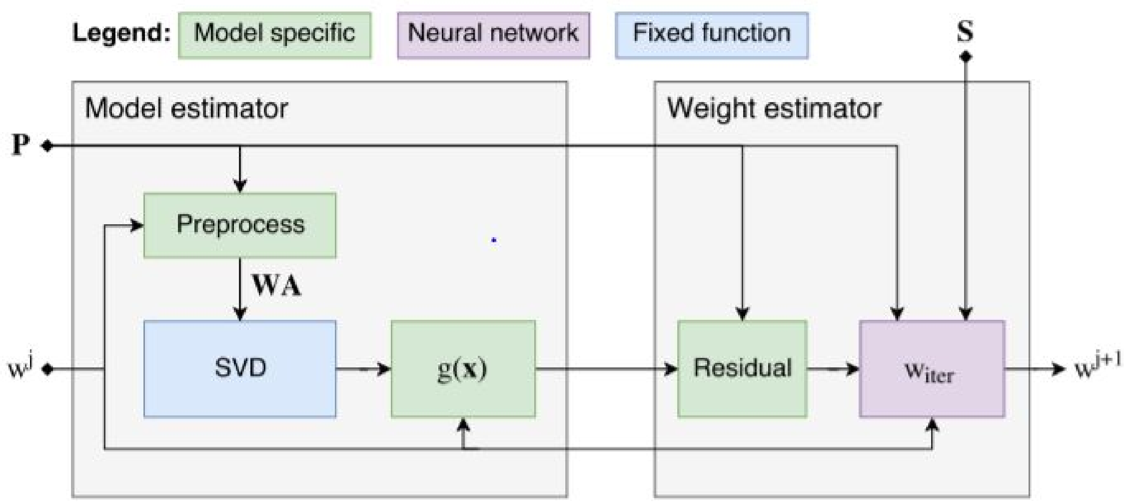
  It was labeled data consisting of 7,200 images of 100 objects. This data was provided to us by the industrial partner which was taken from Columbia University Image Library (COIL-100) software and databases. According to Columbia University, “this dataset is intended for non-commercial research purposes only”.

***Last 4 remaining questions answered in the literature review section*

LITERATURE SURVEY:-

1) PAPER :- Deep Fundamental Matrix Estimation

   RELEVANCE :- This paper presents a technique for the robust calculation of fundamental matrix with deep learning. Calculating a fundamental matrix is key to establishing geometric relationships between two images of the same scene (as, our problem is to estimate the angle of rotation of various objects). The problem is cast as a series of weighted homogeneous least-squares problems, where robust weights are estimated using deep networks.



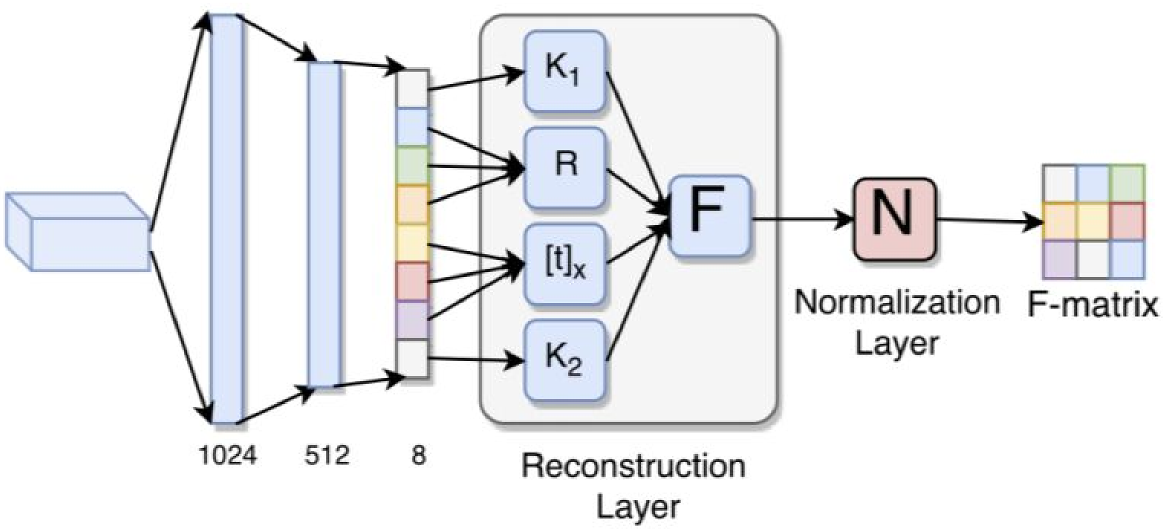
   RESULTS : - The architecture suggested in this paper fits perfectly into standard 3D vision pipelining of Feature extraction, Matching and Model fitting. Experiments on diverse real-world datasets indicate that the presented approach can significantly outperform RANSAC and its variants.

   UNIQUE / DIFFERENCE :- Classical fundamental matrix estimation is done in a precise mathematical fashion, provided that lots of assumptions are being made about the data at hand. This paper suggests that Robust fundamental matrix estimation can be done if the estimator is adapted to the data at hand.

LINK :- http://vladlen.info/papers/deep-fundamental.pdf

2) PAPER :- Deep Fundamental Matrix Estimation without Correspondences

RELEVANCE :- In this paper, an end-to-end trainable convolutional neural networks for F-matrix estimation that do not rely on key-point correspondences, is proposed. Fundamental Matrix is essential to establish relevance between images of the same scene in different angles. Traditional methods rely heavily on the correctness of estimated key-point correspondences, which can be noisy and unreliable. The Architecture below is used for f-matrix estimation.



UNIQUE / DIFFERENCE :- The main problem with fundamental matrix estimation is to preserve it's mathematical properties of having a rank of 2 and with 7-degree freedom. A Reconstruction module and normalized layer is proposed to Address this challenge.

RESULTS :- Performance of the proposed models using various metrics on the KITTI dataset, and show that they achieve competitive performance with traditional methods without the need for extracting correspondences. And In this paper, Two different parameterizations of the F-matrix are considered: one based on the camera parameters, and the other based on the epipolar parametrization.

LINK :-
https://vision.cornell.edu/se3/wp-content/uploads/2018/09/Deep-fundamental-matrix-estimation-camera-ready-2.pdf

3) PAPER :- Finding the exact rotation between two images independently of the translation

RELEVANCE :- For our initial implementation it is important to distinguish the rotation of an object from one image to another, and one major issue we might encounter is the effect of translation of the object. This paper addresses the shortcomings of previous approaches to this issue and proposed a different solution.

RESULT :- The authors discuss using the following approach: First, add gaussian noise. Then find a Gröbner basis. Then select the best solution. Perform model selection to address the problem of converging to 0. And finally applying nonLinear refinement. This results in a very accurate computation of both rotation and translation of an image.

UNIQUE / DIFFERENCE :- This results in a much better computation of angle and translation than other approaches before, including minimizing the sum of squared differences, as well as taking short-term integrals.

LINK :- http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.899.4095&rep=rep1&type=pdf