

CS 726 Project Proposals (Both proposals in this document):

TEAM NAME: MVP-J

Roll no, NAME: (140050015, MOHIT VYAS) and (14B030021, JAYNEEL PAREKH)

PROPOSAL 1

TITLE: Application of Graph CNN's to Image Style Transfer

Key Reference: Unifying local and non-local signal processing with graph CNNs
(<https://arxiv.org/pdf/1702.07759.pdf>)

Problem Statement: We wish to implement the new framework introduced in the work to one of the following specific problems: Image Style Transfer (most probable as of now), Color Transfer, Some new application if we can come up with. (The first two are given in the reference).

Dataset & Evaluation: There are no datasets as they use random shallow networks and possibly because of this, we won't need any additional resources for training. Evaluation is also not quantified and subjective in nature, according to the problem, and can only be judged by observing results on different inputs.

Overall Ideas: They introduce a new framework for graph CNNs, that is CNN for graph signals, which takes two input variables: a signal and a graph structure. They also propose a definition for convolutions on arbitrary graphs, in particular define non-local convolutions. They show 3 different applications of this framework (Style transfer, Color Transfer, Denoising). Image Style Transfer refers to transferring the style (mainly texture at various scales and color palette) of one image, typically a painting, to a different image, typically a photo.

Keywords: Graph signals, CNN, Image style transfer

PROPOSAL 2

TITLE: Human Pose Estimation (2D) using Stacked Hourglass Networks

Key Reference: Stacked hourglass networks for human pose estimation
(<https://arxiv.org/pdf/1603.06937.pdf>)

Problem Statement (as of now): It is to predict the 2D pose of a human in a given image. 2D pose prediction refers to predicting joint locations for a given set of joints in the image co-ordinates.

We are also exploring the option of doing 3D pose prediction (this stacked hourglass network is used as building block for some systems).

Dataset & Evaluation: MPII Human Pose is the most popular dataset for this task and so we wish to use that. Evaluation metrics are explained in the reference (Percentage of correct key-points metric).

Overall Ideas: They define a new deep CNN for predicting joint locations for humans in an image. The basic building block of the network is the stacked hourglass module which integrates (using residual layers), features maps at various scales to capture local and global information. It's implementation is a bit resource intensive, so the extent of our implementation depends upon the resources we can arrange.

Keywords: Stacked Hourglass, 2D/3D Pose prediction, Human Pose prediction