ARIA
APPLIED RESEARCH IN ACTION

SCE-MAE: Selective Correspondence Enhancement with MAE for Self-Supervised Landmark Estimation

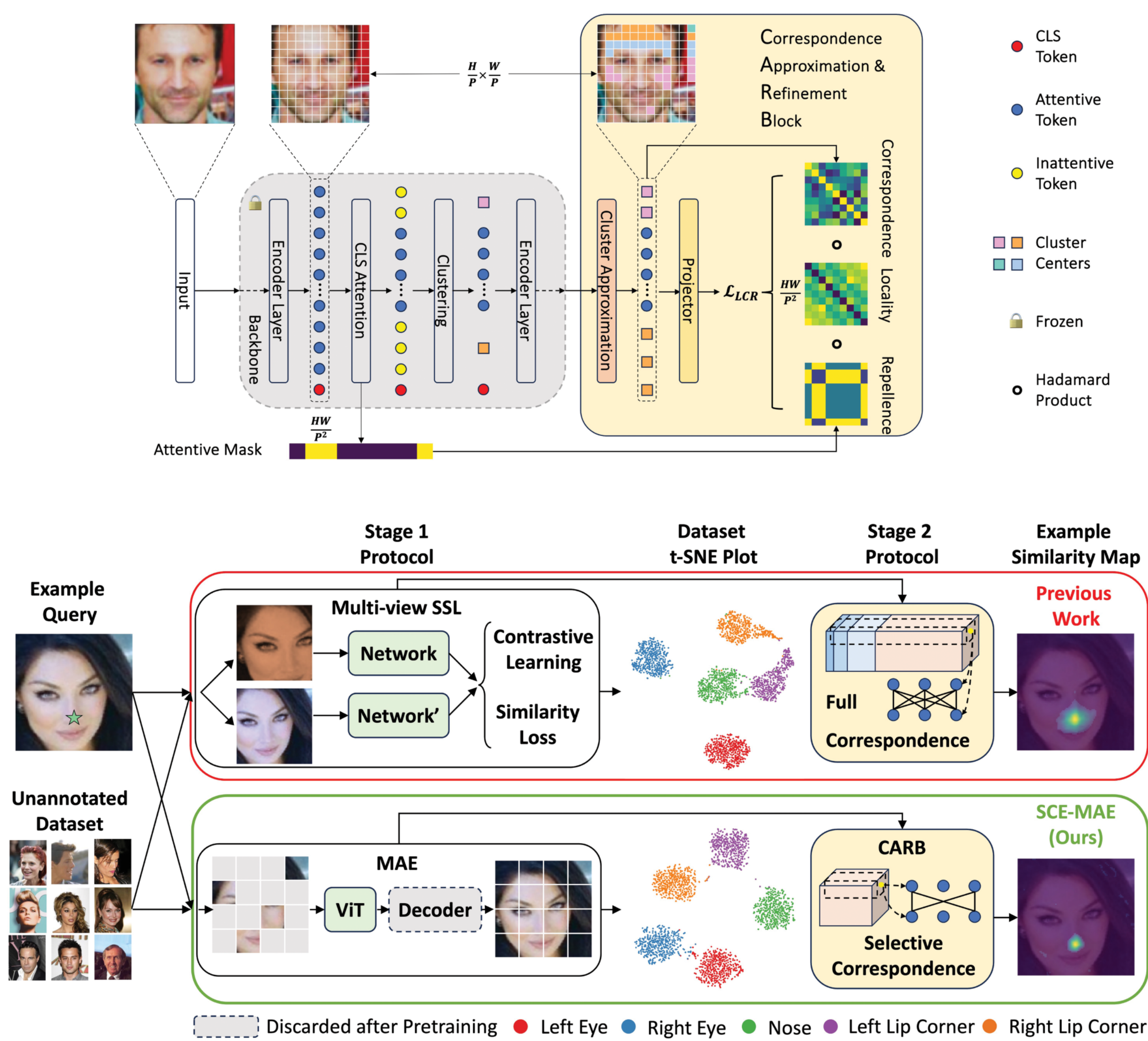## Learning dense discriminative representations from unlabelled data

Kejia Yin

David B. Lindell
ACADEMIC SUPERVISOR

Varshanth Rao
INDUSTRY SUPERVISOR

PROJECT SUMMARY

Although facial landmark detection has been well-studied in supervised setting, its dense and error-prone nature makes the annotation process both expensive and noisy limiting the deployment on many downstream tasks. To address this issue, recent works leveraged self-supervised learning (SSL) to learn meaningful representations from unlabelled data. However, the applied SSL methods only operate their objective on the instance level neglecting the dense prediction nature of the task. In this work, we introduce SCE-MAE, an approach to better exploit face data and learn discriminative dense representations in a self-supervised fashion. We adopt Masked AutoEncoder (MAE) as our first stage SSL protocol and further refine the learned feature with our novel Correspondence Approximation and Refinement Block (CARB) in the second stage. By using a simple density peak clustering algorithm and our proposed Locality-Constrained Repellence loss, CARB hones the selected correspondence between the approximated dense features. Extensive experiments demonstrate the effectiveness and robustness of our method, in which we outperform existing SOTA methods on landmark matching (~20%-48%) and detection (~9%-15%) by a large margin.

MODIFACE

Computer Science
UNIVERSITY OF TORONTO

Master of Science in
Applied Computing