Berrijam Project

Justin Regoli, Keith Yang, Patrick Nasr, Mauricio Diaz Villarreal, Daniel Lu

Overview

Why is image classification relevant in todays society?

Components of ML systems that enable automated visual understanding

Drive innovation in computer vision research
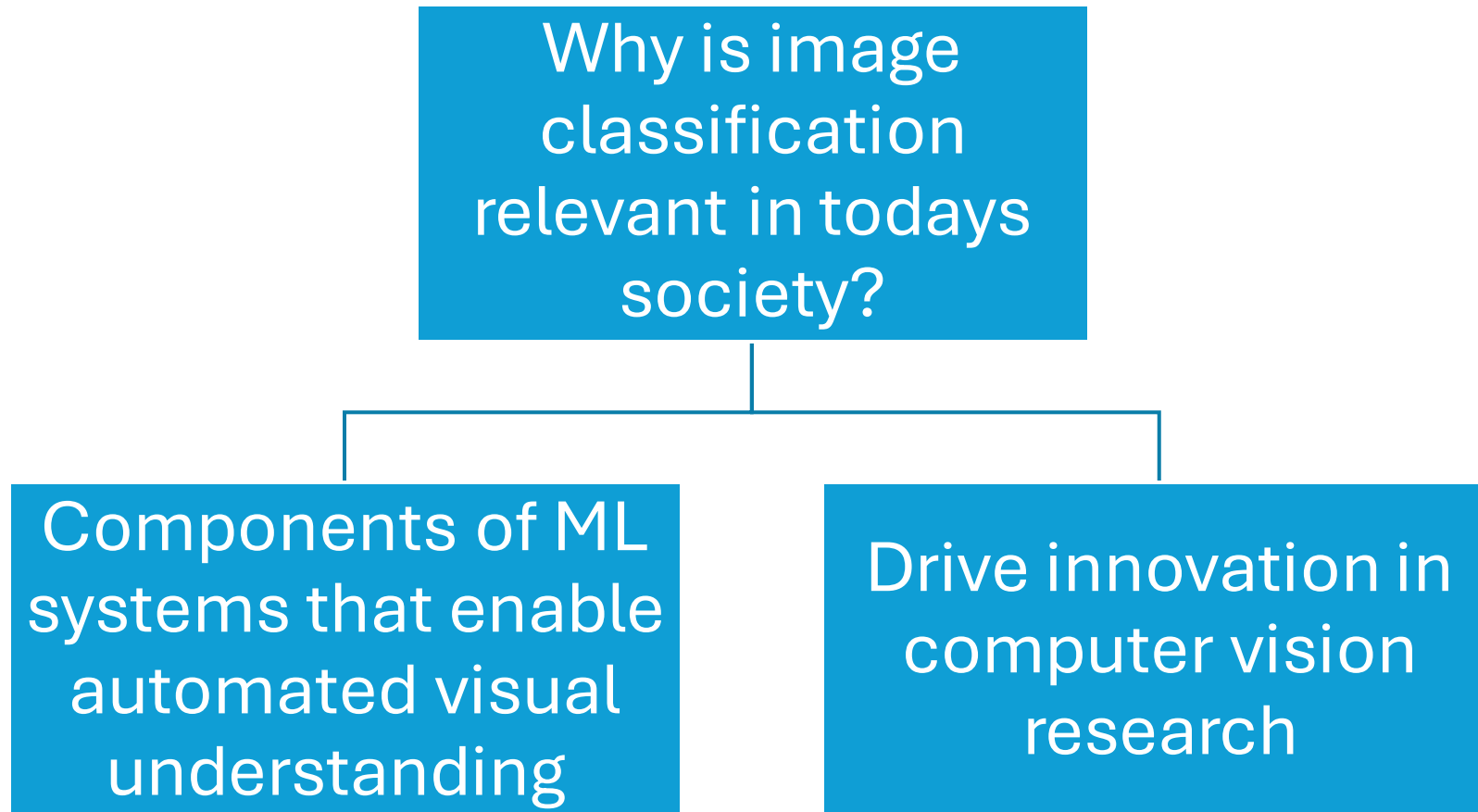
Selection of pretrained models

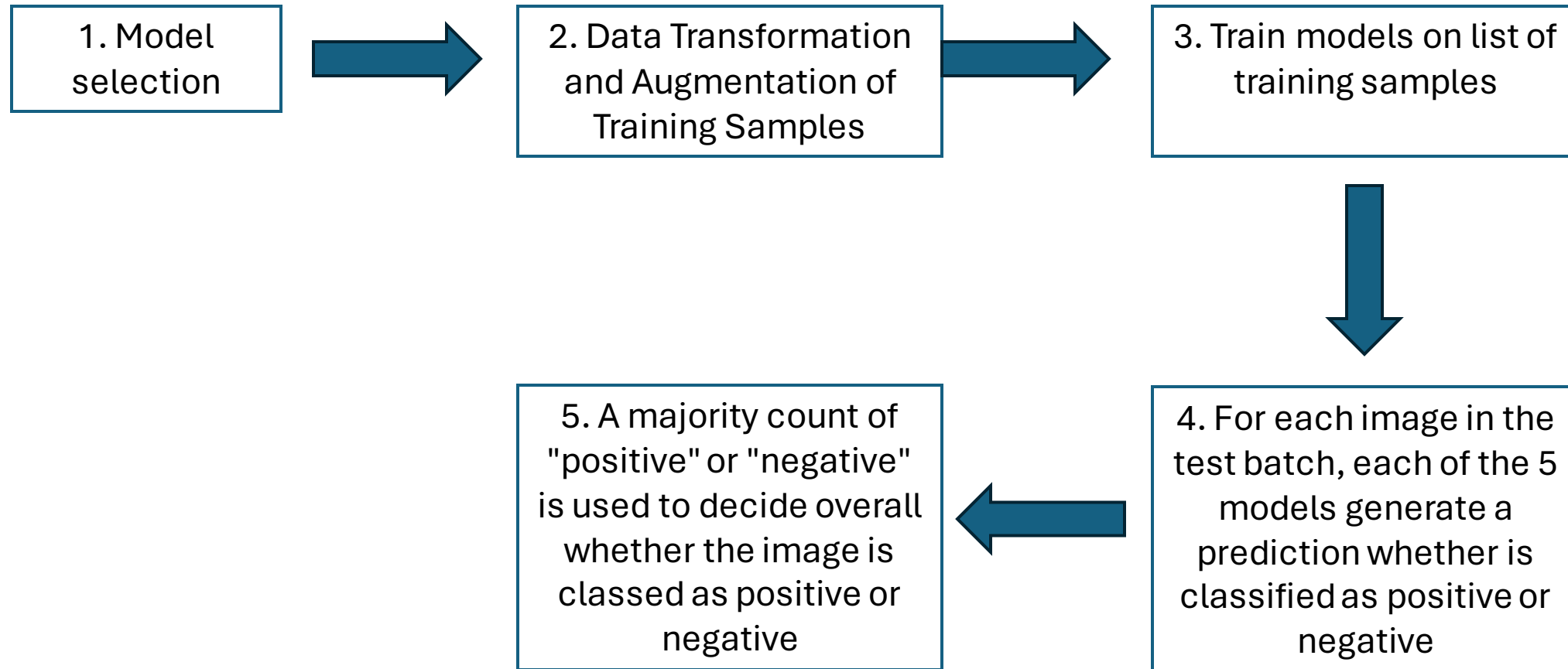Microsoft Swin model: uses window shifting to analyse data.

Google ViT: Treats image recognition like a sequence prediction problem

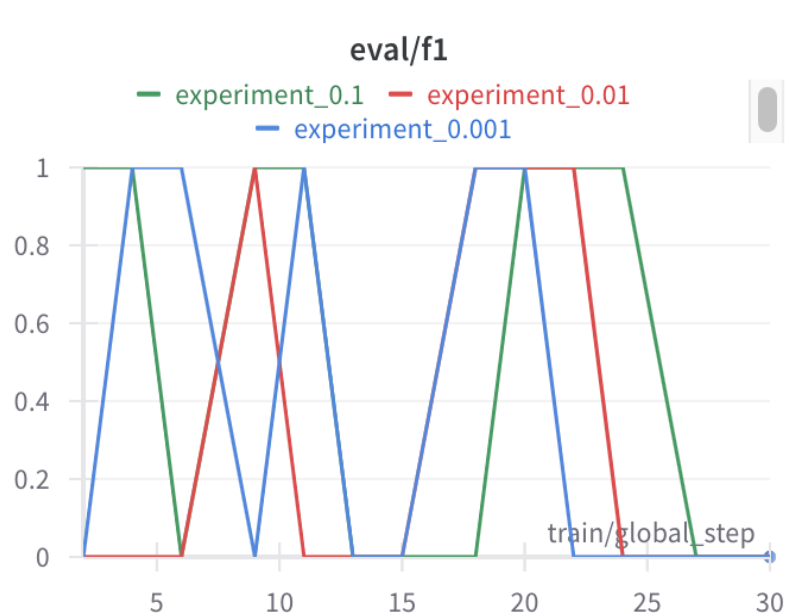Caidas Swin: Enhances resolution of images before applying Swin transformers

Facebook ConvNeXt: Combines a CNN with tranformers
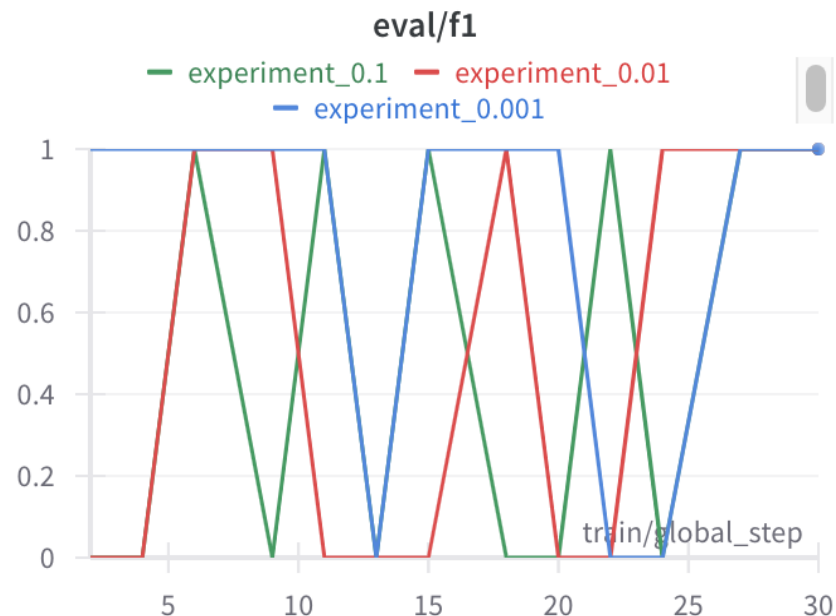
Microsoft Swin model 2: improved version of model #1

Methodology

1. Model selection → 2. Data Transformation and Augmentation of Training Samples → 3. Train models on list of training samples

↓

4. For each image in the test batch, each of the 5 models generate a prediction whether is classified as positive or negative

←

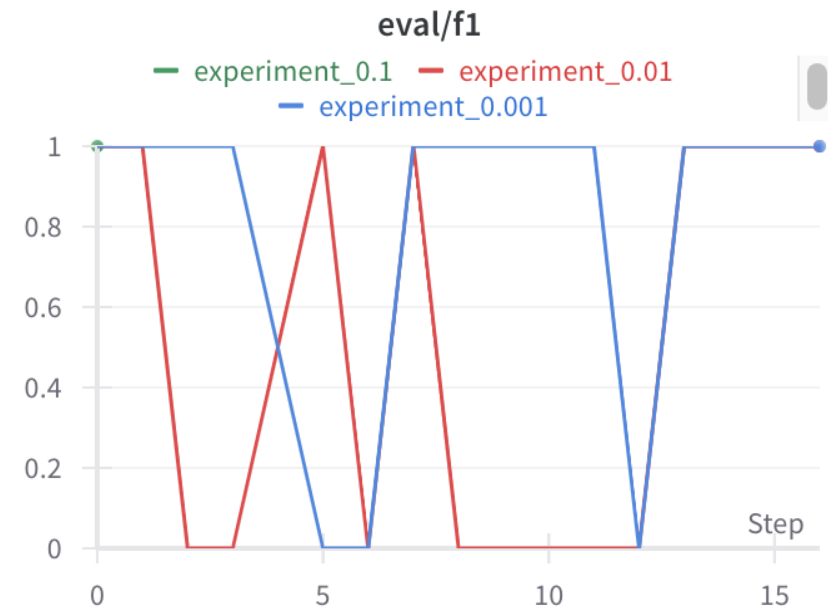5. A majority count of "positive" or "negative" is used to decide overall whether the image is classed as positive or negative

Results and Discussion

ViT

Swin

CNNs