CNN is not all you need-1

- **Improving Machine Vision Using Human Perceptual Representations: The Case of Planar Reflection Symmetry for Object Classification**

  Questions:
  > How closely machine vision representations match human perception ?

  Approach:
  - Compares human and machine using perceptual distances. (called perceived dissimilarity)
  * It is a visual search paradigm

  Ex- A cat and dog may be far in it but a leopard and a jaguar will be nearer.
  on a self calculated distance

  * For Humans
  Formula= 1/Reaction Time prop dissimilarity[derived from inverse]

  Checked the correlation between odd number and even number subjects to check robustness.

  * For machines
  23 vision models and extract features using Gabor Filter,CNNs,SIFT etc.) and check euclidean distances between feature vectors.

  Compare human representations(based on even and odd subjects) based on all 26k objects and find corrected split half correlation between them.(to negate any bias and show robustness) i.e rc

  Compare both human and machine representations based on all 26k objects and find correlation coeff. between them. i.e rm

  Find % variance= (rm/rc)^2

  ML models can only represent 55% variance compared to humans.

  Now ensemble all models and find it against Human perception

and found 68% variance now.(**Answering 1st question**)

> Do machine vision models deviate systematically from human perception ?(Any Pattern or systematic deviation)

* Most objects that have symmetry have underestimated pairs

To quantify it; find variation of residual error(for symmetric objects) between object pairs what have displayed changes.
Plot correlation between residual error and symmetry .

Symmetric objects are distinctive in perception wrt to machine vision(**Answering 2nd question**)

* Most objects that have shared features have overestimated pairs.
To quantify it; find mean residual error(for shared features) between object pairs what have displayed changes.

Also checked for objects with various views and found the same

Objects with share features or parts are similar in perception to most vision models.(**Answering 2nd question**)

> Can one improve machine vision models using human perception ? Only tried to improve for symmetry based object pairs.

Used *late fusion* by introducing a symmetry features extracted on a negative of the image then to a classifier and fused with o/p from normal can on positive image.

then put both probs to a classifier and train a classifier to classify +ve and - ve class.

When we use this symmetry + CNN on imagine we find a 5% increase on some categories on IMAGENET.

This augumentation can be help.

* Why this works ?
last fc layers on CNN are weak for symmetry representations fuse

that with symmetry and its magic.

To quantify we see the most important node in the last year and drop it off(NOT CLEAR)

units that work best for object classification are weaker representations of symmetry

Yes can be improved if augumentation with symmetry info to cnn.

> Limitations:
- Is this result even generalisable ?
- Why only take symmetry. ?