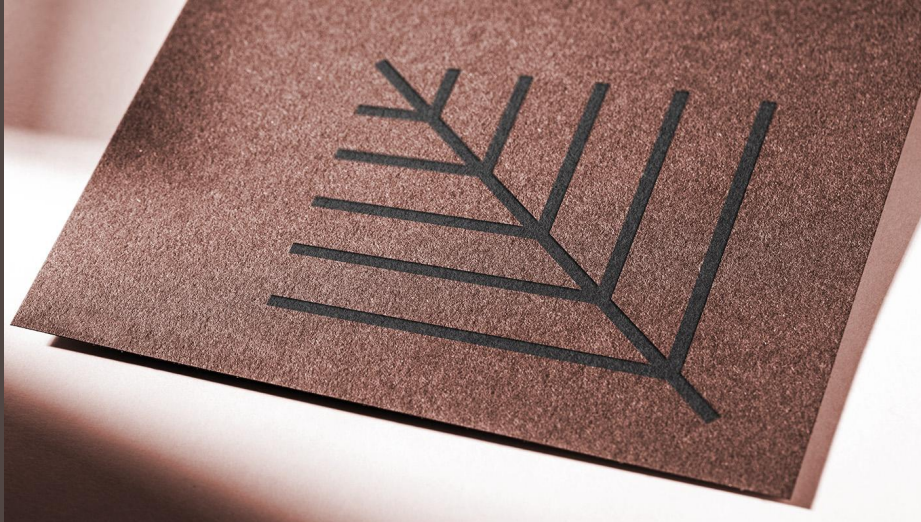


THE PREDATORY WASP OF THE PALISADES

MAN'S QUEST
FOR MEANING
THROUGH THE
EYES OF THE
MACHINE



TABLE OF CONTENTS



1 

PART I

Introduction

2 

PART 2

Algorithms &
Neural
Networks

3 

PART 3

Implementation
& Results

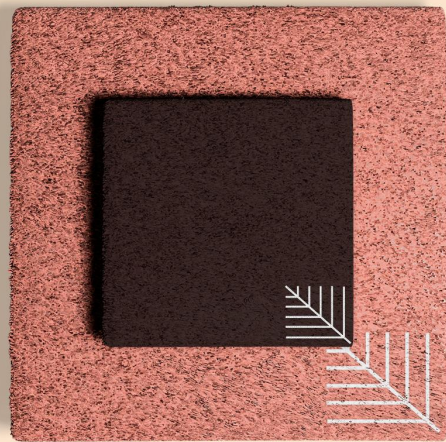
4 

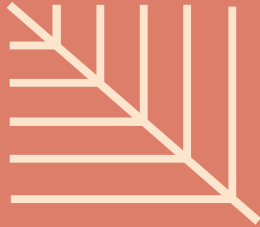
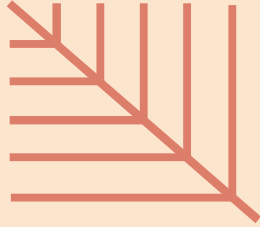
PART 4

Conclusion



INTRODUCTION





**ASSISTED
VISION**

PROBLEM STATEMENT

PROBLEM

Accessibility for
blind and
illiterate



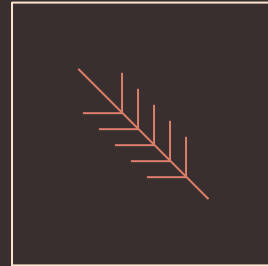
CHALLENGE

Human Like
Descriptions



SOLUTION

Assistive Vision
AI



THE POETRY OF HUMAN DESCRIPTIONS

"Thinking outrageously I write in cursive

I hide in my bed with the lights on the floor

Wearing three layers of coats and leg warmers

I see my own breath on the face of the door

Oh I am not quite sleeping

Oh I am fast in bed

There on the wall in the bedroom creeping

I see a wasp with her wings outstretched

North of Sylvanna we swim in the Palisades

I come out wearing my brothers red hat

There on his shoulder my best friend is bit seven times

He runs washing his face in his hands

Oh how I meant to tease him

Oh how I meant no harm

Touching his back with my hand I kiss him

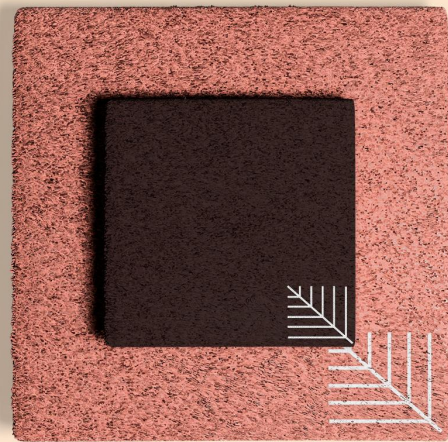
I see the wasp on the length of my arm"



- **SUFJAN STEVENS**

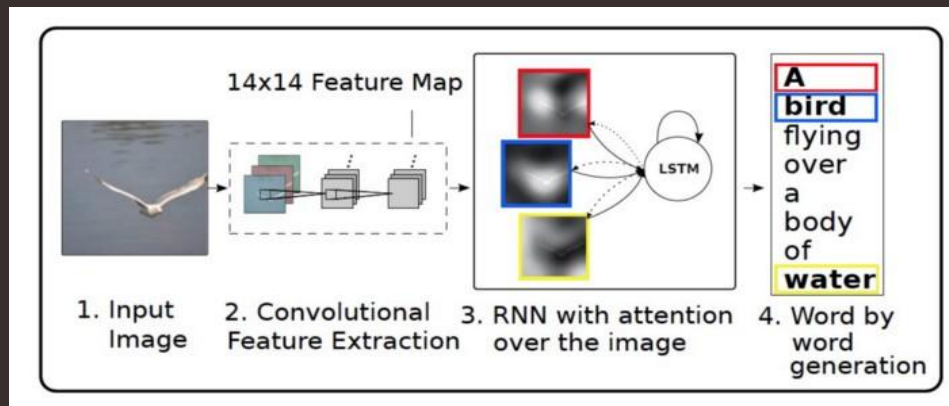
2. 

ALGORITHMS & NEURAL NETWORKS



SHOW ATTEND AND TELL

XU, ET AL



$$p(s_{t,i} = 1 \mid s_{j < t}, \mathbf{a}) = \alpha_{t,i}$$
$$\hat{\mathbf{z}}_t = \sum_i s_{t,i} \mathbf{a}_i.$$

PAPER HIGHLIGHTS

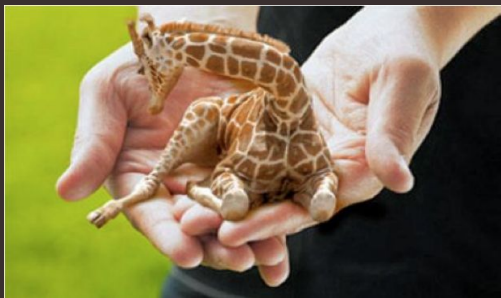
Attention learning over caption learning

Word by word generation

Stochastic gradient descent using adaptive learning rate algorithms

SELF CRITICAL SEQUENCE TRAINING FOR IMAGE CAPTIONING

RENNIE, ET AL



PAPER HIGHLIGHTS

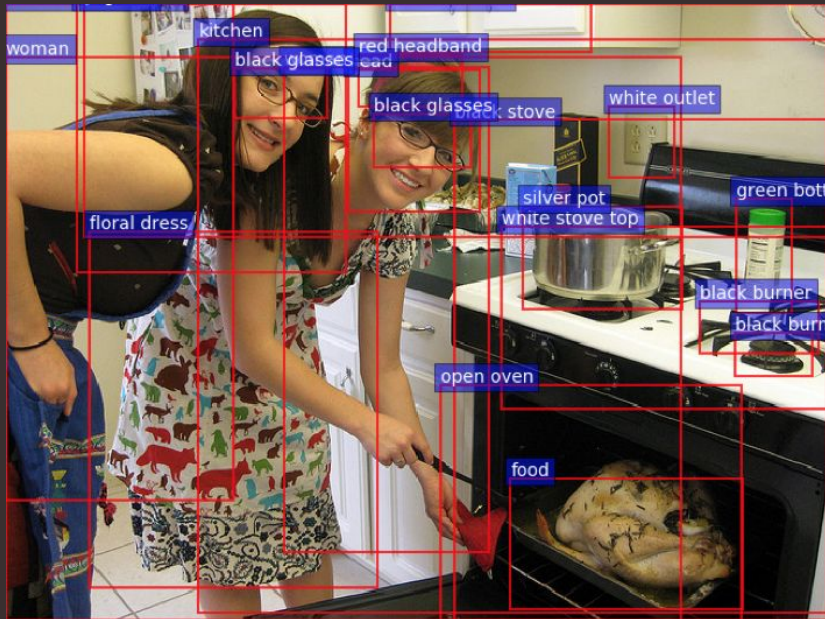
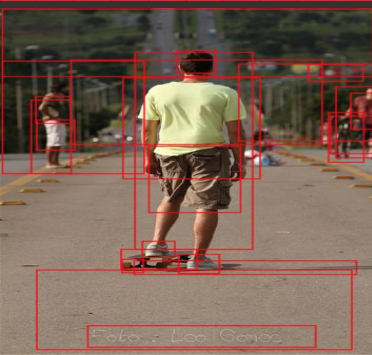
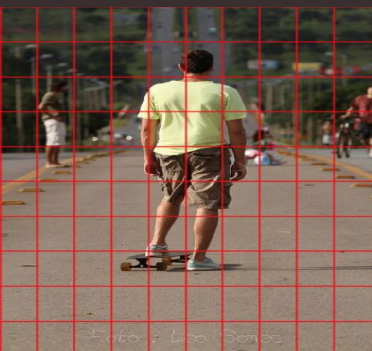
Optimizing image captioning systems using reinforcement learning

Policy Gradient with REINFORCE

Self-critical sequence training (SCST)

BOTTOM-UP AND TOP-DOWN ATTENTION FOR IMAGE CAPTIONING AND VISUAL QUESTION ANSWERING

ANDERSON, ET AL



PAPER HIGHLIGHTS

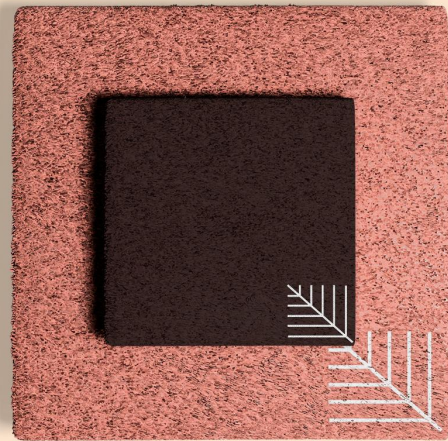
More closely resembles human vision through bottom up and top down

Faster R-CNN

Simplest Implementation

3. 

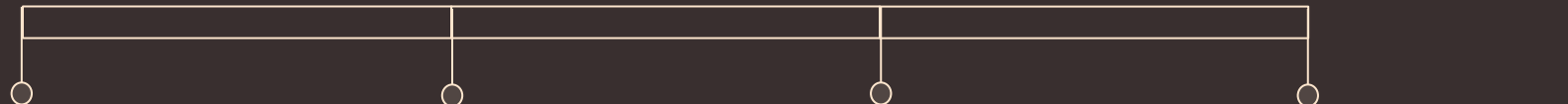
IMPLEMENTATION & RESULTS



158,000 IMAGES
790,000 CAPTIONS

100,000+ Epochs
(And Counting!)

SHOW ATTEND AND TELL MODEL



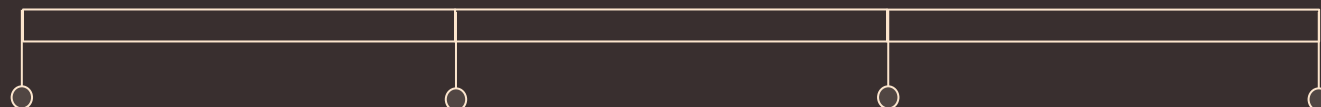
Training Time	Library Used	Dataset(s) Used	Highest Score
5 Full days	Tensorflow Keras	MSCOCO Flickr30k Flickr8k	BLEU-4: 14.87



EXAMPLE I.

<start>
a red car in
the forest
<end>

SELF CRITICAL SEQUENCE TRAINING MODEL



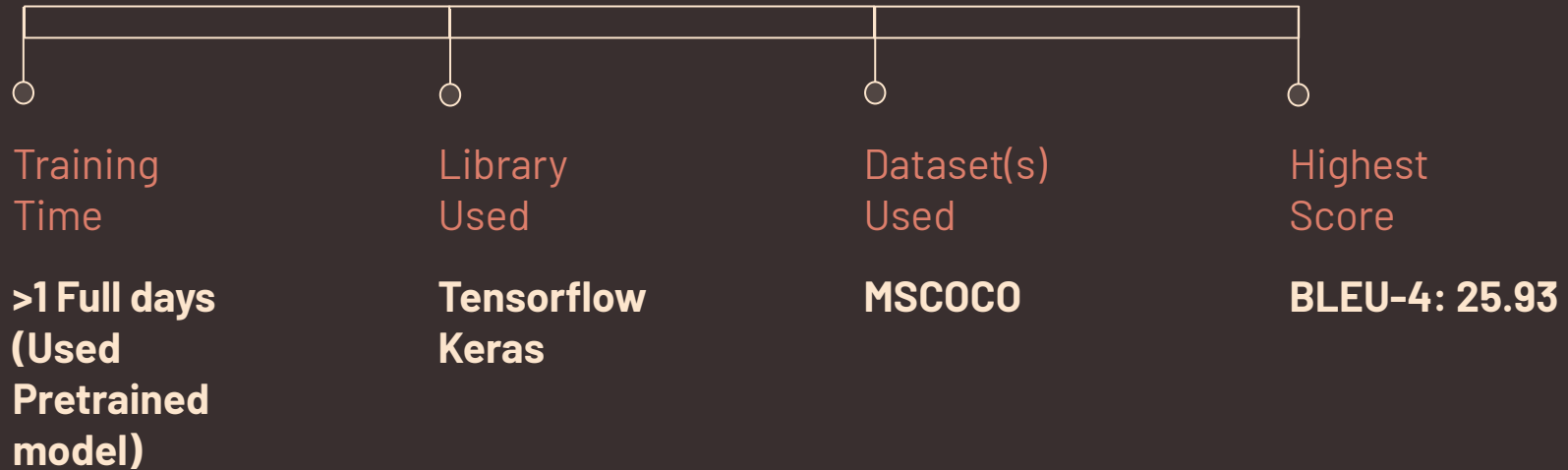
Training Time	Library Used	Dataset(s) Used	Highest Score
3 Full days (Ongoing)	Pytorch	MSCOCO	BLEU-4: 19.08



EXAMPLE 2.

<start>
a clock on a
brown
building
<end>

BOTTOM-UP TOP-DOWN MODEL



The diagram illustrates a Bottom-Up Top-Down Model structure. It features a horizontal line at the top with four vertical lines extending downwards from it, each ending in a small circle. These circles are positioned above the column headers of a table below. The table has four columns: Training Time, Library Used, Dataset(s) Used, and Highest Score. Each column contains a specific value or metric.

Training Time	Library Used	Dataset(s) Used	Highest Score
>1 Full days (Used Pretrained model)	Tensorflow Keras	MSCOCO	BLEU-4: 25.93



EXAMPLE 3.

<start>
a group of
men on skis
in the snow
<end>

SCORE COMPARISONS



SHOW & TELL

SCST

**BOTTOM-UP
TOP-DOWN**

**OFFICIAL
SCORES**

25.0

35.04

36.90

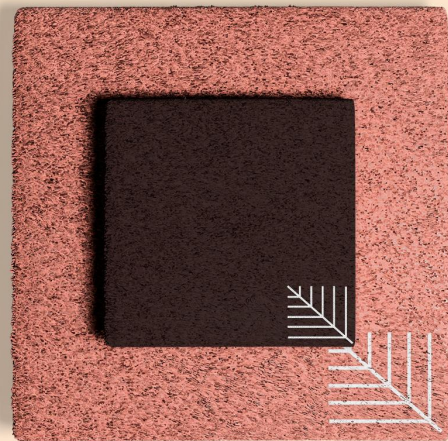
**ASHANTI
SCORES**

14.87

19.08

25.93

4. CONCLUSIONS



CONCLUSIONS

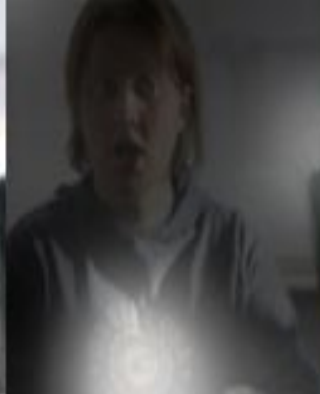
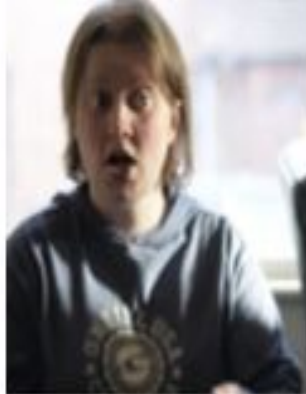
AND TAKEAWAYS

- We are nowhere near fully human level descriptiveness with Machine Learning
- Current implementations can only operate circumstantially
- Ethical issues could range from misclassification of individuals, to objects
- The work is being done to address these issues, it will just require time and seemingly unlimited resources.
- At this point in time humans are far superior in describing our natural world, and perhaps it should stay that way.





A large white bird standing in a forest.



A woman holding a clock in her hand.



A man wearing a hat and
a hat on a skateboard.



A person is standing on a beach
with a surfboard.



A woman is sitting at a table
with a large pizza.



A man is talking on his cell phone
while another man watches.

THANK YOU!



QUESTIONS?

ASHANTI.O.JABRI@GMAIL.COM | [LINKEDIN.COM/ASHANTIJABRI](https://www.linkedin.com/in/ASHANTIJABRI)