Visual Question and Answering on Blur Images

PRESENTATION

By: Group No 34

Aayush (2020009)
Anwar (PhD22009)
Adarsh (PhD22504)
Deepak (PhD22023)
Sahil (2020115)
Sibin (2020307)

Problem Statement

- The problem addressed in this research paper is the need for a reliable method to determine the recognizability of blurry images.
- The research aims to develop a threshold-based approach that can accurately determine the recognizability of blurry images.
- The proposed method will be evaluated using a comprehensive dataset of blurry images.
- the performance will be compared to existing methods to demonstrate its effectiveness.

Motivation

- Blurred images can pose a significant challenge for image recognition systems used in various applications, including medical diagnosis, security systems, and autonomous vehicles.
- This problem is particularly concerning for visually impaired individuals who rely on image recognition systems to navigate their surroundings.
- our work aims to develop a method to set a threshold for recognizability to help image recognition systems determine whether an image is recognisable.
- This threshold can improve the performance of image recognition systems on blurry images,

Literature Review

Main categories of method reviewed.

1. Image Sharpness Measures
   - Traditional methods for blur Detection
   - Sharp images typically have high contrast edges, while blurred images have smoother transitions

2. Deep Learning based methods
   - More recent approach for blur detection
   - Leverages power of CNNs
   - CNNs are type of deep neural network designed for image processing

Image sharpness measures

1. Gradient based measures:

   Compute the magnitude of the gradient of an image, which is a measure of the rate of change of the image intensity. A sharp image will have a high magnitude gradient, while a blurred image will have a low magnitude gradient

2. Laplacian-Based Measures:

   compute the Laplacian of an image, which is a measure of the second derivative of the image intensity. A sharp image will have a high Laplacian, while a blurred image will have a low Laplacian. Common Laplacian-based measures include the variance of the Laplacian and the Laplacian of Gaussian (LoG) operator.

3. Frequency-Based Measures:

   Frequency-based measures analyze the frequency content of an image using Fourier or wavelet transforms. Sharp images will have high-frequency content, while blurred images will have low-frequency content. Common frequency-based measures include the power spectrum, the normalized Laplacian pyramid, and the wavelet-based blur detection method

Deep learning based methods

1. Classification-Based Methods:

   Classification-based methods use a CNN to classify an image as either sharp or blurred. The CNN is trained on a dataset of labeled images, where each image is labeled as either sharp or blurred. During training, the CNN learns to extract features from the image that are useful for distinguishing between sharp and blurred images. Once the CNN is trained, it can be used to classify new images as either sharp or blurred.

2. Regression-Based Methods:

   Regression-based methods use a CNN to predict a blur score for an image. The blur score is a continuous value that represents the degree of blur in the image. The CNN is trained on a dataset of labeled images, where each image is assigned a blur score based on some metric of image quality, such as the variance of the Laplacian. During training, the CNN learns to extract features from the image that are useful for predicting the blur score. Once the CNN is trained, it can be used to predict the blur score for new images

3. Generative Adversarial Networks (GANs):

   GANs are a type of deep learning model that consists of two networks: a generator network and a discriminator network. The generator network learns to generate realistic images, while the discriminator network learns to distinguish between real and generated images. GANs have been used for image restoration tasks, such as deblurring and super-resolution. In the context of blur detection, a GAN can be trained to generate realistic blurred images, which can then be used to train a CNN to distinguish between sharp and blurred images.

Novelty

1) Blurry images are inherently more difficult to analyze and interpret.
2) Blurry images can result from
    a) camera shake,
    b) low light conditions
    c) poor focusing.
3) As a result, the objects and details in the image may be unclear, making it challenging for the algorithm to accurately identify and classify them.
4) VQA on blurry images has a number of potential applications
    a) Surveillance (blurry footage)
    b) Security
    c) Medical imaging (blurry images due to low radiation)
5) It is important to define how much blurriness can be handled by VQA system

Methodology

1. Dataset:

   The vizwiz task has a large collection of images out of which many images are blurred because these images/photos are taken by blind people.

2. Pre-processing:

   Identify images which look blurry for further processing.

3. Feature Extraction:

   We use image feature detection algorithms to extract visible features from the image.

4. Answerability detection:

   Based on the detected features we decide if the image is suitable for question answering or not.

5. Threshold setting:

   Based on the percentage blurriness of various images and their answerability we derive a threshold which can be used as reference for declaring an image as answerable or not.

Results

Recognisability/Blur detection tool (Focus measure):

71 lines (71 sloc) | 2.07 KB

Raw Blame

In [1]:

```python
import cv2
import os
import csv

def variance_of_laplacian(image):
    return cv2.Laplacian(image, cv2.CV_64F).var()

# Path to directory containing images
image_dir = 'fmap/detectron/train'

# Create CSV file and write header row
with open('blur_detection_results_train.csv', 'w', newline='') as file:
    writer = csv.writer(file)
    writer.writerow(['filename', 'blur_score', 'blur_detection_result'])

    # Loop through all images in directory
    for filename in os.listdir(image_dir):
        if filename.endswith('.jpg') or filename.endswith('.png'):
            # Read image and convert to grayscale
            image = cv2.imread(os.path.join(image_dir, filename))
            gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)

            # Calculate blur score and determine blur detection result
            blur_score = variance_of_laplacian(gray)
            if blur_score < 100:
                blur_detection_result = 'blurry'
            else:
                blur_detection_result = 'not blurry'

            # Write row to CSV file
            writer.writerow([filename, blur_score, blur_detection_result])
```

Output range - [0,20395.00187]

(Focus measure $\propto$ image quality)

Results

Classification using Xception Model

Code: Link

We executed the code using two sets of data from the main dataset:

1) Train_set1

   Output: clothes, clothes, cardboard, shoes, clothes, glass, shoes, glass, shoes, plastic

   Accuracy according to human verification was 33%

2) Train_set2

   Output: clothes, shoes, clothes, shoes, clothes, shoes, plastic, clothes, shoes, plastic

   Accuracy according to human verification was 40%

Conclusion

- We proposed a novel threshold-based approach to accurately determine the recognizability of blurry images.
- Our proposed method of visual question and answering on blur images is a novel and challenging task that can have significant implications for the field of computer vision and information retrieval.
- Our methodology included data collection, pre-processing, and feature extraction, which are critical steps in visual question and answering on blur images.
- Ultimately, the outcome of this work is the definition of a threshold for blurred images