REAL TIME FACE RECOGNITION USING INTEL NEURAL COMPUTE STICK

Shashank Nigam, Ramdas KrishnaKumar, Padmini Ramesh,Srikant,Abhineet

Institute of Systems Science, National University of Singapore, Singapore 119615

ABSTRACT

Face recognition finds a variety of application from phone unlocking to person identification in different scenarios. While face detection can be efficiently implemented through existing algorithms such as Viola-Jones[1], face recognition is computational expensive.This reduces efficiency of real time face recognition on edge devices like CCTV deployed in real time monitoring. Siamese Networks for one shot face recognition[2], can be used to recognize faces with just one sample image.Edge devices have low resources for computation, storage and power. AI based applications are computation intensive and power consuming. Intel's Movidius Neural Compute Stick 2 (NCS)[3], is a plug and play AI device with 16 cores providing high CPU acceleration for edge devices. This ensures scalability of AI based application at low power, low resources and high computation. Deploying Siamese Neural network on NCS we implement an efficient real time face detection on edge devices.

Index Terms— Siamese Neural Network, Movidius Neural Compute Stick, Voila-Jones, Edge Computing, face Recognition

1. INTRODUCTION

Real time face recognition requires a robust mechanism and low latency to recognize identity of human face. In modern face recognition system the conventional pipeline consist of four stages detect -¿ align-¿represent-¿classify. Depending on the algorithm face recognition after detection can involve different methods for feature extraction and classification. Some of the prominent feature extraction methods may involve [4]

- Geometric detection: It looks at different landmark features such as relative position, size, and/or shape of the eyes,nose,cheek and jaw
- 3-Dimensional recognition
- Skin Texture Analysis: Turning unique lines,patterns and spots apparent in a person's skin into a mathematical space
- Thermal Camera: Detecting human face using thermal imaging.Thereby ignoring the different accessories such as glasses, hats or makeup.

landmark detection e.g relative position, size, and/or shape of the eyes,nose,cheek and jaw [4].

Although most of the approaches are highly accurate in recognizing human face, they however require complex expensive computation and expensive sensors. Traditional methods are not scalable and imposes high resource and computational cost on the edge device.Most of the traditional Edge device like Camera,CCTV used mostly for surveillance, have limitation on resources such as memory, computation, and power. Traditional face recognition system requires high computation, high power and high resource cost, this further involves computing on the cloud thereby increasing the latency.

2. PROPOSED APPROACH

For real time face recognition on edge devices following steps can be taken into consideration

- Face detection using Voila-Jones algorithm [1]
- Face recognition using Siamese Network [2]

2.1. Siamese Network

One Shot face recognition using Siamese neural network [2]. Each of the face has unique feature which can be encoded using a convolution neural network. Different images of the same person tend to have similar feature. Thus the distance between the two encoded features of the same person would be less in comparison with distance of the encoded features of different person. Below Network architecture is employed for training the network

The model is trained with set of two images at a time. Each of the set of 2 images can belong to the same person or different person. The model is further optimized over contrastive loss[5].

$$L(W,(Y,\vec{X_1},\vec{X_2})^i)=(1-Y)(D_w)^2+(Y)(max(0,m-D_w))^2 \quad (1)$$

Where X_1 and X_2 define the input images, while m determines the margin between the similar and dissimilar images. Once the network is trained we would require a database of only single images to recognize faces.

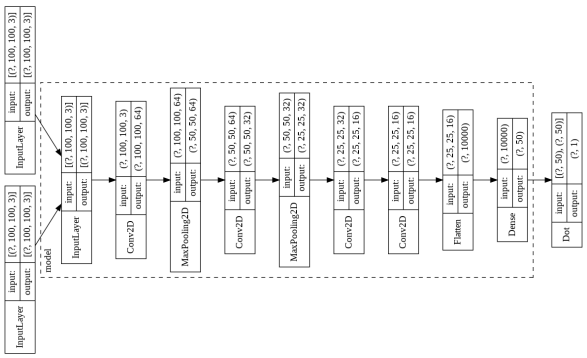

Fig. 1. Siamese Network

2.2. Model Conversion

Edge devices are used for capturing data on a daily basis.In order to make infrastructure operate efficiently we would require the a device which shares major computation load and at the same time consumes Low power. We make use of Intel Neural Compute stick 2 (NCS) for the purpose.NCS is a USB, low power (8W), fan-less, high computational (16 cores) device [3]. This makes real time inferencing possible on edge devices. Deploying the trained network on NCS requires the model to be optimized for the device. The model optimization is done using Open-Vino and involves following steps Freezing the Tensorflow model. This converts the variables to constant weights Optimizing the model using Open-Vino. This step converts the frozen model into two inference model XML and BIN, which is understood by the NCS. Since NCS can operate on only floating point 16 (FP16) the model weights are converted into FP16 at this stage. This can be further summarized as below

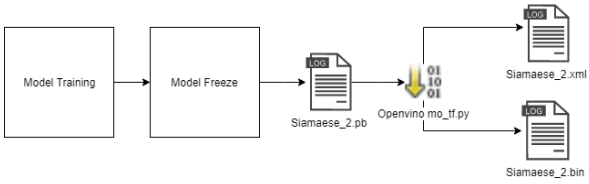

Fig. 2. Model Conversion

2.3. Real-time face recognition

Once the model is converted in appropriate inference files it can be further used for deploying into NCS. At production we maintain a small set of database containing single image of the possible candidates to be recognized. The data obtained from camera can be used for face detection using Voila-Jones. The detected segments are further passed to Siamese Network for face recognition deployed on NCS device. The architecture can be summarized as below

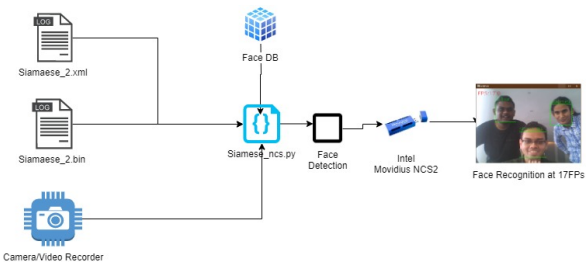

Fig. 3. Face Recognition

3. EXPERIMENTAL RESULTS

We tested the application running on CPU of Intel i7 2nd Generation and NCS Below are the results for the same.

Table 1. The performance comparison.

Device	Person Recognition(FPS)
Intel i7 2nd Gen	35 fps
Movidius Neural Compute Stick 2	33.5 fps

4. CONCLUSIONS

Intel Movidius Neural compute Stick 2 can give a comparable performance with CPU in low power which makes it apt for deploying it on an edge device. With proper trained Siamese Network the architecture can be used for real time face recognition.

5. REFERENCES

[1] Michael Jones Paul Viola, "Rapid object detection using a boosted cascade of simple features," in *CVPR,2001*, 2001.

[2] Yann LeCun Roopak Shah Jane Bromley, Isabelle Guyon, "Signature verification using a siamese time delay neural network," in *NIPS'93*, 1994, pp. 737–744.

[3] "Intel movidius neural compute stick 2:https://software.intel.com/en-us/neural-compute-stick," .

[4] "Face recognition:https://en.wikipedia.org/wiki/facial$_recognition_system$

[5] "Dimensionality reduction by learning an invariant mapping," 2006, CVPR,2006.