International Journal of Engineering & Technology, 7 (3) (2018) 1208-1215

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET
doi: 10.14419/ijet.v7i3.13169
Research paper

GPU Accelerated real-time Melanoma Detection

Ajai Sunny Joseph [1] *, Elizabeth Isaac [2]

[1] P.G Student, Mar Athanasius College of Engineering, Kothamangalam, Kerala, India
[2] Assistant Professor, Mar Athanasius College of Engineering, Kothamangalam, Kerala, India
**Corresponding author E-mail: ajay.sunny@gmail.com*

Abstract

Melanoma is recognized as one of the most dangerous type of skin cancer. A novel method to detect melanoma in real time with the help of Graphical Processing Unit (GPU) is proposed. Existing systems can process medical images and perform a diagnosis based on Image Processing technique and Artificial Intelligence. They are also able to perform video processing with the help of large hardware resources at the backend. This incurs significantly higher costs and space and are complex by both software and hardware. Graphical Processing Units have high processing capabilities compared to a Central Processing Unit of a system. Various approaches were used for implementing real time detection of Melanoma. The results and analysis based on various approaches and the best approach based on our study is discussed in this work. A performance analysis for the approaches on the basis of CPU and GPU environment is also discussed. The proposed system will perform real-time analysis of live medical video data and performs diagnosis. The system when implemented yielded an accuracy of 90.133% which is comparable to existing systems.

Keywords: *Cancer Detection; GPU; Melanoma; Skin Cancer; Deep Learning*

1. Introduction

The early diagnosis and detection of diseases with the help of computer-aided systems is getting more prominence every day. It is like at someday in future, computer systems may take over this area reducing the workload of a doctor. Existing systems can process medical images and perform a diagnosis based on Image Processing technique and Artificial Intelligence. Such systems perform processing with the help of large complex systems at the backend. This results in a higher cost for data processing. Graphics Processing Unit (GPU) is a new technology capable for finding out solutions to the computational problems in all the engineering and medical fields. In the medical industry, this supports the processing of higher dimensional data. GPU computation has provided a huge edge over the central processing units (CPU) with respect to computation speed. It is highly parallel, multithreaded, consists of multiple core processors and has high memory bandwidth to give the solution to the computational problems. A reason for the evolution of powerful GPUs are the constant demand for greater realism in computer applications. During the past few decades, the computational performance of GPUs has increased much more quickly than that of conventional CPUs. Hence, It plays a major role in the field of modern industrial research and development. GPU has already achieved a significant speed (twox-1000x) than CPU implementation on various fields.

Cancer in its different varieties is becoming one of the most popular diseases across the world. Early detection and diagnosis is crucial in case of survival for this type of diseases. A variety of machine learning tools including Artificial Neural Networks (ANNs), Bayesian Networks (BNs), Support Vector Machines (SVMs) and Decision Trees (DTs) have been widely applied in cancer research for the development of predictive models, resulting in effective and accurate decision making. Human malignant melanoma is a highly metastatic cancer type that is substantially resistant to standard therapeutic modalities. Although the survival rate of melanoma patients has been improved over the past decades melanoma risk and overall mortality escalate yearly. As per a survey conducted by World Health Organization (WHO) [1], it is estimated that between 2 and 3 million non-melanoma skin cancers and 132,000 melanoma skin cancers occur globally each year. The studies done by American Cancer Society (ACS) [2] predicts that about 91,270 new melanomas will be diagnosed in 2018 at USA (about 55,150 in men and 36,120 in women). It also states that about 9,320 people of these are expected to die of melanoma (about 5,990 men and 3,330 women). These rates have been rising for the last 30 years and this shows the importance of addressing this issue. Diagnosing melanoma can be a complex task. An article in the Journal of Investigative Dermatology [3], explains the complexity of the clinical process of diagnosing melanoma, it involves several skin examination components such as anamnestic data (medical history), comparative and differential recognition as well as pattern analysis. The manual detection can be a lengthy process and it requires highly trained dermatologists. The accuracy of dermatologists is estimated to be about 75 to 85%. Most of these deaths due to melanoma are occurring because of the failure to detect it on the right time. If melanoma is recognized and treated early, it is almost always curable, but if it is not, the cancer can advance and spread to other parts of the body, where it becomes hard to treat and can be fatal.

GPUs together with deep learning approaches can be of great help in developing a real time detection system for these kind of diseases. The detection of these diseases, as said earlier, is done by a system with the help of machine learning tools. A GPUs contribution comes here as it saves on the training time of these machine learning tools by significant margin as compared to normal systems running on CPUs. Deep Learning plays a vital role in the early detection of cancer. This type of diseases requires detection and treatment at right time. It is also very difficult to detect cancerous growths at their earlier stages by human observation. As per the

studies conducted by NVIDIA [4] deep learning helped to reduce the error rate on breast cancer detection by 85%. Fortunately, deep learning has shown capabilities in achieving higher diagnostic accuracy results in comparison to many domain experts. While this may be an issue of contention with physicians, for many would-be victims, the technology cannot come soon enough. Hierarchical or structured deep learning is a modern branch of machine learning that was inspired by the human brain. This technique has been developed based on complicated algorithms that model high-level features and extract those abstractions from data by using neural network architecture that is similar but much more complicated.
The existing systems [5-9] identifies diseases from images. The proposed work aims at identifying diseases from live footages. The use of GPU accelerated computing techniques based on Compute Unified Device Architecture (CUDA) and TensorFlow to speed up the processing is proposed. A machine learning framework, which will be fine tuned to analyze the frames of video, based on Convolutional Neural Network (CNN) is also proposed in this work. The objective of the proposed work is to develop a real time system that takes live footage as input, then analyze this footage to detect any malicious tumorous or cancerous growths, then identify and indicate the affected path, then finally identify and provide the type and severity of disease.

2. Materials and methods

Most of the melanomas on first look resembles very close to that of normal skin pigmentations or pimples. This makes it really confusing and hence sometimes it is very difficult to identify the skin pigmentation in our body as malign or benign. Using an automated system that will closely observe different cases and learn from them will help for early diagnosis and accurate detection of melanoma.
A most commonly used clinical method for distinguishing melanoma is the 'ABCDE' method [10] shown in Figure 1. The ABCDE criteria were intended as a simple tool that could be implemented in daily life, a mnemonic " as easy as ABC " to alert both the layperson and primary care physician to the clinical features of melanoma. It was introduced in 1985 as a method for diagnosing melanoma. ABCDE acronym stands for Asymmetry, Border Irregularity, Color, Diameter and Evolving. When it was initially introduced, it was only ABCD. Later 'E' was added into the acronym in 2004. These parameters have been widely used by Dermatologists to diagnose melanoma. Each criteria has certain features that are recognized to classify melanoma into benign and malignant.
This method is a generally applicable and precise method. But there are some special cases where this method also fails. Such as, If the melanoma is having a very small diameter then it violates the conditions about melanoma stated as per this method. Also there are cases where birth marks are misidentified as melanoma as per the guidelines of this method. So this means although these kind of methods have improved easiness in differentiating malign and benign moles. But there are still chances for wrong diagnosis of moles.

Deep Learning frameworks based on CNNs can be used to find a solution to this misclassification issue. Automated systems for the real time detection of melanoma can be implemented using CNNs. Significant performance improvisations can also be done by using GPUs, as CNNs are highly GPU supported. CNNs are able to extract the non-human extractable and very complex features from its input. They are also capable of learning all the details without engineering the features.

The architecture of the proposed system is shown Figure 2. The system takes live video footages or real time video captured by a camera device as input.

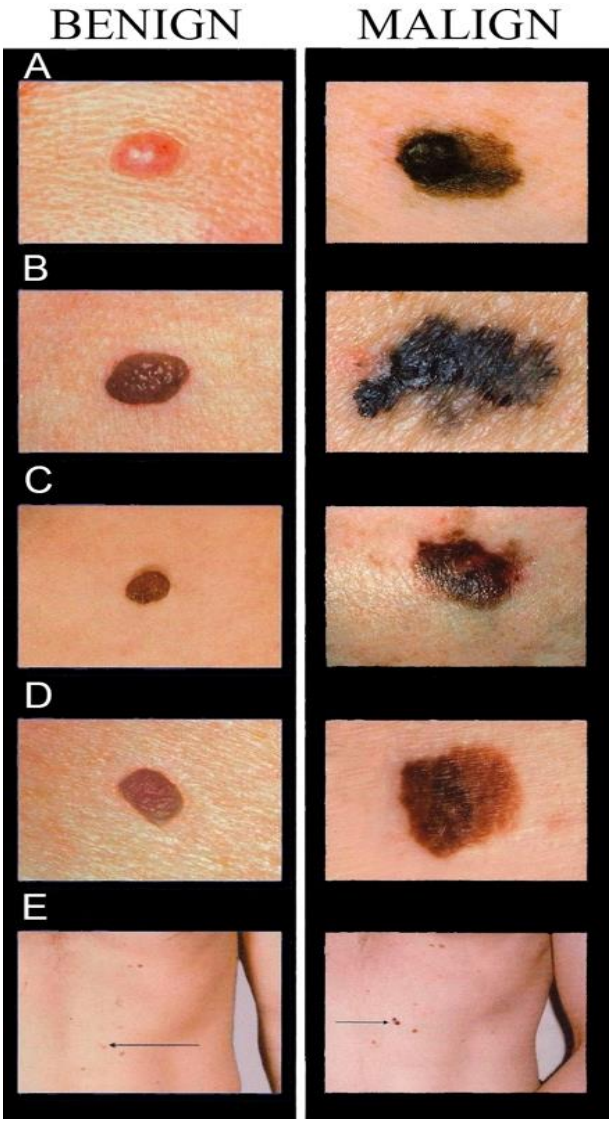

Fig. 1: Benign v/s Malign Pigmentations Based on ABCDE Method where A Stands for Asymmetry, B for Border Irregularity, C for Color, D for Diameter and E for Evolving.

This live footages is analyzed for the detection of melanoma using deep learning framework based on CNN. It does this through a combination of filters using machine learning, image recognition and extraction of global and local features.

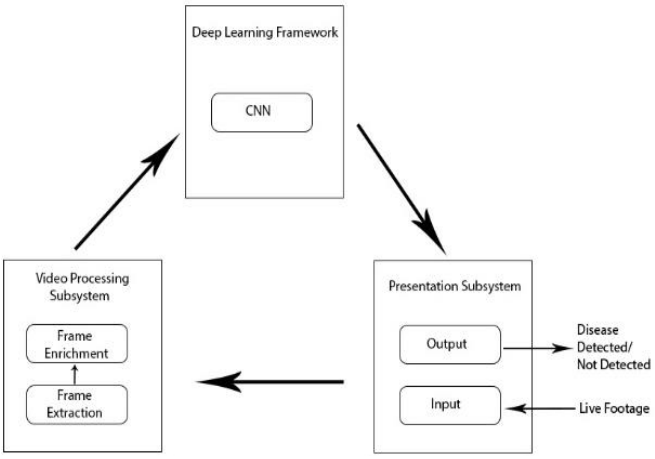

Fig. 2: The Proposed System Architecture. It Consists of Three Subsystems Presentation Subsystem, Video Processing Subsystem and A Deep Learning Framework.

3. Proposed system

The proposed system is divided mainly into three subsystems,

3.1. Presentation subsystem

This subsystem captures the real-time frames and passes it to other subsystems for processing. It also presents the output of the real-time analysis to the medical examiner. The most challenging aspect here is that the presentation should not introduce any delays, which would make the system unsuitable for live examinations.

3.2. Video processing subsystem

Video processing subsystem is responsible for performing the frame extraction from the input footage. It is responsible for handling the enhancement and processing. This step includes enrichment of the frame by removing noises, improving the contrast etc. and thus improving the overall quality. This subsystem makes use of OpenCV 3.0 and FFMPEG 3.4.2 libraries for implementing frame extraction and enrichment process.

3.3. Deep learning framework

The deep learning framework consists of convolutional neural network. Regular neural networks do not scale well into images. But convolutional neural network architectures make the explicit assumption that the inputs are images, which allows us to encode certain properties into the architecture. The proposed system is implemented with three different types of convolutional networks and performance is analysed for them.

3.4.1. U-net

U-net [11] is a fully convolutional neural network. It consists of 23 convolutional layers. The main idea is to supplement a usual contracting network by successive layers, where pooling operators are replaced by upsampling operators. Hence, these layers increase the resolution of the output. The architecture of U-net is shown in Figure 3. The left side path is the contracting path and right side path is the expansion path. Two 3x3 unpadded convolutions are repeatedly applied and each of them is followed by a Rectified Linear Units (ReLU) and a 2x2 max pooling operation with stride 2 for downsampling. The number of feature channels are doubled in each downsampling. At each step in expansion path, Upsampling of feature map is followed by a 2x2 convolution (up-conv) which halves the number of feature channels, a concatenation with the correspondingly cropped feature map from the contracting path, and two 3x3 convolutions, each followed by a ReLU. Cropping is performed because of the loss of border pixels in every convolution. The final layer is a 1x1 convolution which maps 64 component feature vector to desired classes.

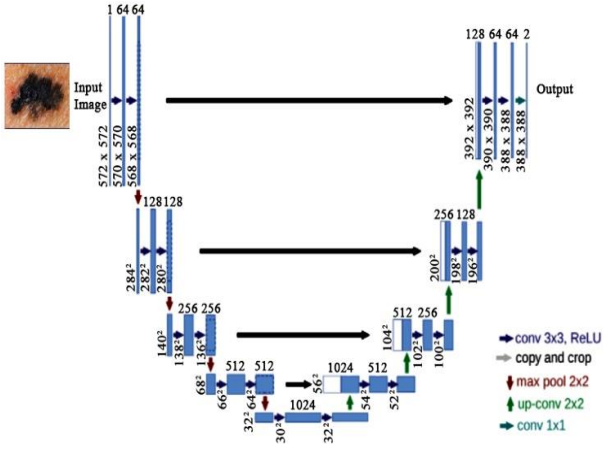

Fig. 3: Architecture of U-Net Consisting of 23 Convolutional Layers.

During the training phase, benign and malign images and their corresponding segmentation maps were used for training the network. Soft-max for the network is defined as,

$$p_k(x) = \exp(a_k(x))/\left(\sum_{k'=1}^{K} \exp(a_{k'}(x))\right) \quad (1)$$

Where $a_k(x)$ denotes the activation in feature channel k at the pixel position $x \in \Omega$ with $\Omega \subset Z^2$. K is the number of classes and $p_k(x)$ is the approximated maximum-function. I.e. $p_k(x) \approx 1$ for the k that has the maximum activation $a_k(x)$ and $p_k(x) \approx 0$ for all other k. The cross entropy then penalizes at each position the deviation of $p_{l(x)}(x)$ from 1 using,

$$E = \sum_{x \in \Omega} w(x) \log(p_{l(x)}(x)) \quad (2)$$

Where $l : \Omega \rightarrow \{1, \ldots, K\}$ is the true label of each pixel and $w : \Omega \rightarrow R$ is a weight map that we introduced to give some pixels more importance in the training. The separation border is computed using morphological operations. The weight map is then computed as,

$$w(x) = w_c(x) + w_0 . \exp\left(-\frac{(d_1(x)+d_2(x))^2}{2\sigma^2}\right) \quad (3)$$

where $w_c : \Omega \rightarrow R$ is the weight map to balance the class frequencies, $d_1 : \Omega \rightarrow R$ denotes the distance to the border of the nearest cell and $d_2 : \Omega \rightarrow R$ the distance to the border of the second nearest cell. In our experiments we set $w_0 = 10$ and $\sigma \approx 5$ pixels.

3.4.2. ResNet

ResNet is short for Residual Neural Networks, which were introduced in late 2015 by Microsoft. It is a type of Convolutional Neural Network used for image recognition. It consists of hundreds of layers. The main disadvantage of deep convolutional neural networks such as U-net is the degradation problem. Degradation problem is caused due to the increase in the depth of network; as a result the accuracy gets saturated and then degrades rapidly. This kind of degradation is not caused by overfitting further increase in layers will result in increase in training error. ResNets address this problem by the introduction of identity shortcut connections or identity mapping as shown in Figure 4. Normal CNNs will be having direct mapping, where the output of a layer is F(x), instead of using identity mapping, where the output of a layer is given as $F(x) + x$. Here x denotes the identity mapper value. A trend observed in ResNet as mentioned by He et al.[12] is that as it goes deeper or increasing number of layers is actually reducing the error rate as compared to plain networks where it happens the opposite. So learning is better in RCNN which has more stacked layers as compared to others. Also the performance is better with the use of powerful GPU and more data.
The building block of a ResNet is mathematically represented as,

$$y = F(x, \{W_i\})f(x) \quad (4)$$

where x and y denote the corresponding input and output of the layers. $F(x, \{W_i\})$ denotes the residual mapping. For the example shown in Figure 4. $F = W_2\sigma(W_1 x)$ in which σ denotes ReLU.

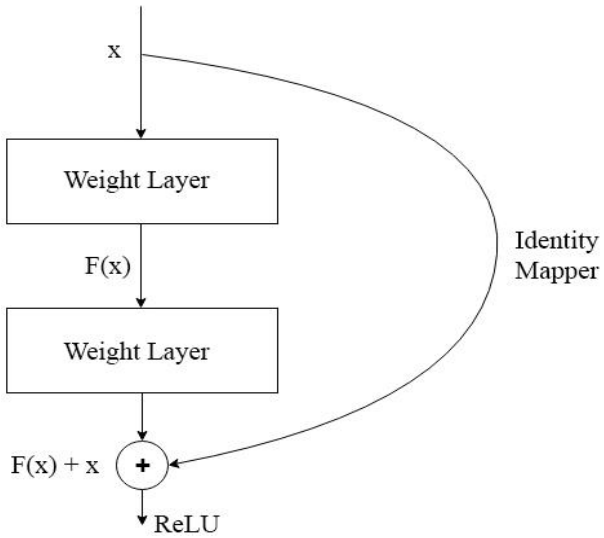

Fig. 4: The Building Block of A Resnet, Where and Denote the Corresponding Input and Output of the Layers. Relu Stands for Rectifier Linear Units. F(X) is the Output of the Weighted Layer and X Denotes the Identity Mapper.

The proposed ResNet is shown in Figure 5. It consists of 5 set of convolution layers. The detailed specification is as shown in Table 1. The convolutional units in these [5] layers together contribute 101 layers.

Table 1: Resnet 101 Specification

Layer Name	Output Size	No. of Units
Convolution Layer 1	112x112	[7x7, stride 2, channel 64] x 1
Convolution Layer 2	56x56	[3x3 max pool, stride 2] x 1 [1x1, channel 64] x 3 [3x3, channel 64] x 3 [1x1, channel 256] x 3
Convolution Layer 3	28x28	[1x1, channel 128] x 4 [3x3, channel 128] x 4 [1x1, channel 512] x 4
Convolution Layer 4	14x14	[1x1, channel 256] x 23 [3x3, channel 256] x 23 [1x1, channel 1024] x 23
Convolution Layer 5	7x7	[1x1, channel 512] x 3 [3x3, channel 512] x 3 [1x1, channel 2048] x 3
Output	1x1	Average pool, 1000d-fc, soft-back

In this proposed system, the live medical footage from a capturing device, which is a part of the presentation subsystem, interacts with the video processing subsystem. The frame extraction module of the subsystem is responsible for extracting each and every frame from the given input footage. The frame extraction module then passes the captured frame to frame enrichment module which performs image denoising and enhancement. This is done in order to improve the quality of input data into the deep learning framework. If there is a large amount of noise or error in the input data then there are chances that the prediction may be inaccurate which will make the system unreliable. In the next stage the captured frame is passed onto the deep learning subsystem where a trained neural work, which is a Convolutional Neural Network (CNN) proposed in this method, will process the image and performs classification to see whether there are any malign moles exists in the frame and if present, localizes and presents it to the medical examiner who is provided with an output indicating the affected region.

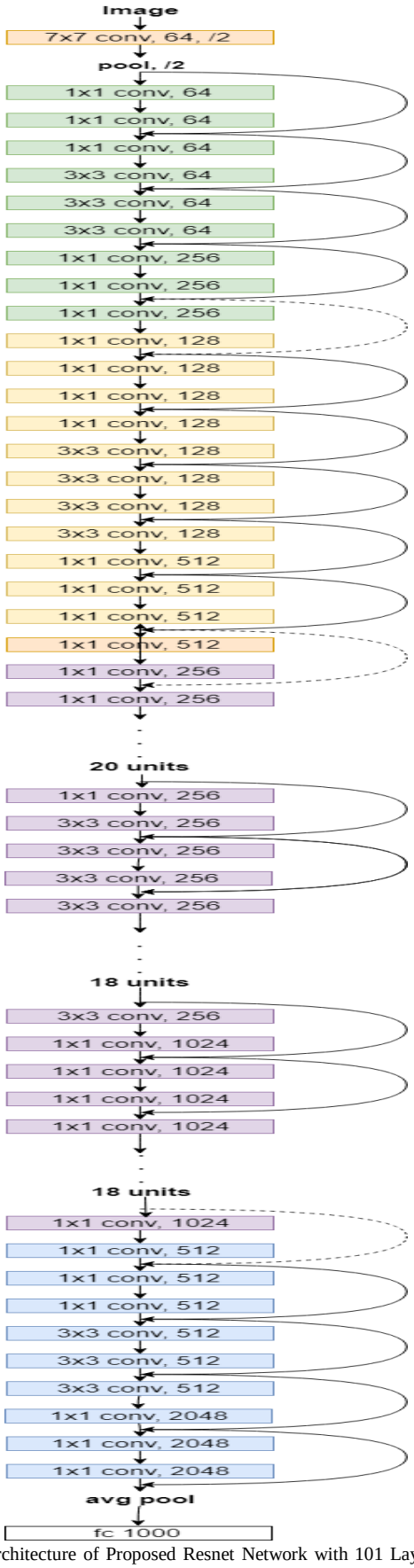

Fig. 5: Architecture of Proposed Resnet Network with 101 Layers. Each Block Denotes the Convolution Operation and Channel Width Associated. The Network Is Having Number Residual Connections. the Convolution Layers Are Divided Into 5 Sections as Indicated with Different Colours. the First Block Consists of Only A Single Unit. Second Block Consists of 9 Units, Third Block Has 12 Units, Fourth Block Has 69 Units and Fifth Block Consists of 9 Units. The Detailed Specifications of Each Block is as Shown in Table 1.

3.4.3. SSD mobile net

The structure of SSD MobileNet[27] is built on depthwise seperable convolutions except for the first layer. All layers are followed by a batchnorm [26] and ReLU nonlinearity with the exception of the final fully connected layer which has no nonlinearity and feeds into a softmax layer for classification. Figure 6 contrasts a layer with regular convolutions, batchnorm and ReLU nonlinearity to the factorized layer with depthwise convolution, 1 × 1 pointwise convolution as well as batchnorm and ReLU after each convolutional layer. Down sampling is handled with strided convolution in the depthwise convolutions as well as in the first layer. A final average pooling reduces the spatial resolution to 1 before the fully connected layer. Counting depthwise and pointwise convolutions as separate layers, MobileNet has 28 layers. The architecture specification is given in Table 2.

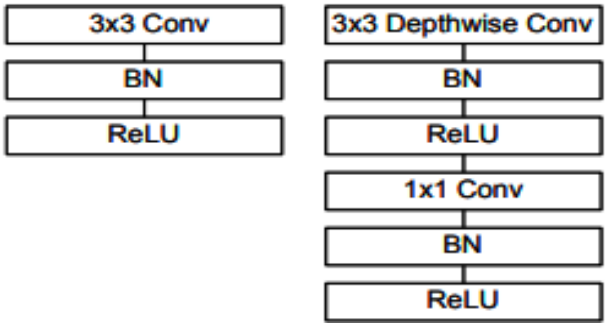

Fig. 6: This Figure Contrasts A Layer with Regular Convolutions, Batchnorm and Relu Nonlinearity to the Factorized Layer with Depthwise Convolution, 1 × 1 Pointwise Convolution as well as Batchnorm and Relu After Each Convolutional Layer. Left: Standard Convolutional Layer with Batchnorm and Relu. Right: Depthwise Separable Convolutions with Depthwise and Pointwise Layers Followed By Batchnorm and Relu.

It is not enough to simply define networks in terms of a small number of Mult-Adds. It is also important to make sure these operations can be efficiently implementable. For instance unstructured sparse matrix operations are not typically faster than dense matrix operations until a very high level of sparsity. Our model structure puts nearly all of the computation into dense 1 × 1 convolutions. This can be implemented with highly optimized general matrix multiply (GEMM) functions. Often convolutions are implemented by a GEMM but require an initial reordering in memory called im2col in order to map it to a GEMM. One×1 convolutions do not require this reordering in memory and can be implemented directly with GEMM, which is one of the most optimized numerical linear algebra algorithms. Nearly all of the additional parameters are in the fully connected layer.

Table 2: SSD Mobile net Specification

Type / Stride	Filter Shape	Input Size
Conv / s2	3 × 3 × 3 × 32	224 × 224 × 3
Conv dw / s1	3 × 3 × 32 dw	112 × 112 × 32
Conv / s1	1 × 1 × 32 × 64	112 × 112 × 32
Conv dw / s2	3 × 3 × 64 dw	112 × 112 × 64
Conv / s1	1 × 1 × 64 × 128	56 × 56 × 64
Conv dw / s1	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 128	56 × 56 × 128
Conv dw / s2	3 × 3 × 128 dw	56 × 56 × 128
Conv / s1	1 × 1 × 128 × 256	28 × 28 × 128
Conv dw / s1	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 256	28 × 28 × 256
Conv dw / s2	3 × 3 × 256 dw	28 × 28 × 256
Conv / s1	1 × 1 × 256 × 512	14 × 14 × 256
5 x Conv dw / s1	3 × 3 × 512 dw	14 × 14 × 512
5 x Conv / s1	1 × 1 × 512 × 512	14 × 14 × 512
Conv dw / s2	3 × 3 × 512 dw	14 × 14 × 512
Conv / s1	1 × 1 × 512 × 1024	7 × 7 × 512
Conv dw / s2	3 × 3 × 1024 dw	7 × 7 × 1024
Conv / s1	1 × 1 × 1024 × 1024	7 × 7 × 1024
Avg Pool / s1	Pool 7 × 7	7 × 7 × 1024
FC / s1	1024 × 1000	1 × 1 × 1024
Softmax / s1	Classifier	1 × 1 × 1000

All of these subsystems has to correlate and work in real-time to produce the output. The deep learning framework is proposed to be accelerated with the help of a GPU. A dominant trend to speed up processing of CPU-intensive tasks is to offload processing tasks to GPUs. Utilizing a GPU and program it using either CUDA will be the right way to achieve real-time performance.

4. Results and discussions

The dataset used by the system is obtained from the International Skin Image Collaboration (ISIC) archive. 956 images of both benign and malign cases are collected from the archive. These images are used for training the system. Original sizes of images vary from 1022×767 to 4288×2848 pixels. Clinical examination videos related to melanoma are used for testing purposes. The system was tested on 10 persons each with melanoma and without melanoma to ensure the realtime detection capability of the system. For other evaluation purposes such as comparison of deep learning models 20 videos were used for testing with each of one having 16 to 30 instances of melanoma in each of them. Tests were conducted on systems with U-net, ResNet, SSD MobileNet and also in a Non-GPU and GPU accelerated environment. The GPU Accelerated environment included TensorFlow for GPU, OpenCV 3.0, Python 3.6 and NVIDIA CUDA Toolkit 9.0 along with an NVIDIA GeForce GTX 1080 Titan Series 4GB DDR5 GPU and Intel Core i7 7th Generation CPU on an ASUS ROG Strix Laptop with 8GB RAM. The Non-GPU environment was setup on TensorFlow, OpenCV 3.0 and Python 3.6 running on an Apple Macbook with 8GB RAM and an onboard Intel HD Graphics 6000 chipset.

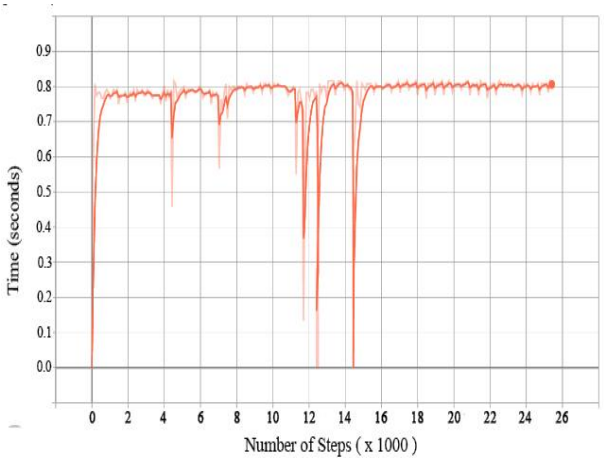

Fig. 7: This Figure Shows the Graph, Which Denotes the Time Taken for Each Step of the Training to Complete in A GPU Accelerated Environment. The X-Axis Denotes the Number of Steps and the Y-Axis Denotes The Time Taken for Each Step to Complete.

Figure 7 denotes the graph showing number of step per second in a GPU accelerated environment. It is evident from the graph that each step takes less than 0.8 seconds in a GPU accelerated environment whereas compared to a CPU environment as shown in Figure 8, each step is taking around 42 seconds to complete.
In GPU accelerated system. In every minute, approximately ≥ 75 steps are completed whereas in CPU environment, it is only 2.333 steps per minute. So the GPU Accelerated system is about 32.147 times faster than the system in CPU environment.

Table 3: No of Malign and Benign Instances in Each Test Input Videos

Video Sample	No off. Malign Instances	No off. Benign Instances
Sample 1	26	4
Sample 2	22	6
Sample 3	22	2
Sample 4	11	7
Sample 5	18	12
Sample 6	13	8
Sample 7	15	8
Sample 8	17	11
Sample 9	9	9
Sample 10	11	11
Sample 11	14	7
Sample 12	20	0
Sample 13	21	5
Sample 14	19	6
Sample 15	0	16

For the real-time testing, 20 specimens were used in which 10 were malign and 10 were benign. For the benign samples, one of them was misclassified and rest of 9 were correctly classified. For the malign samples, only one of them was misclassified and rest of 9 were correctly classified. The sensitivity and specifity of the system were calculated as 0.84 and .98 respectively. The system is giving an accuracy of 90.133%, which is a similar accuracy to that of existing methods. However, the system is expected to give a better accuracy when trained with a larger dataset and tested with better test samples. The quality of the test samples were not the best and was only satisfactory. However, the proposed system is about 32 times faster than existing methods.

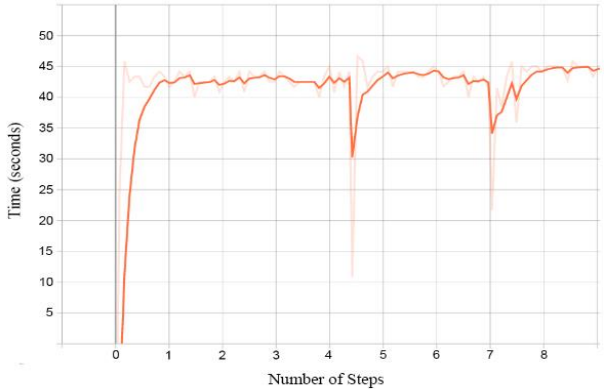

Fig. 8: This Figure Shows the Graph, Which Denotes the Time Taken for Each Step of the Training to Complete in A CPU Only Environment. the X-Axis Denotes the Number of Steps and the Y-Axis Denotes the Time Taken for Each Step to Complete.

The system is implemented with three deep CNNs U-net, ResNet and SSD MobileNet. The comparison of their performance is shown in Figure 8. The accuracy calculated for each video samples run on both systems is shown below on Figure 9. The system implemented with ResNet obtained an accuracy of 90.133 and outperformed the systems implemented on U-net and MobileNet having an accuracy of 89.066% and 89.434% respectively.

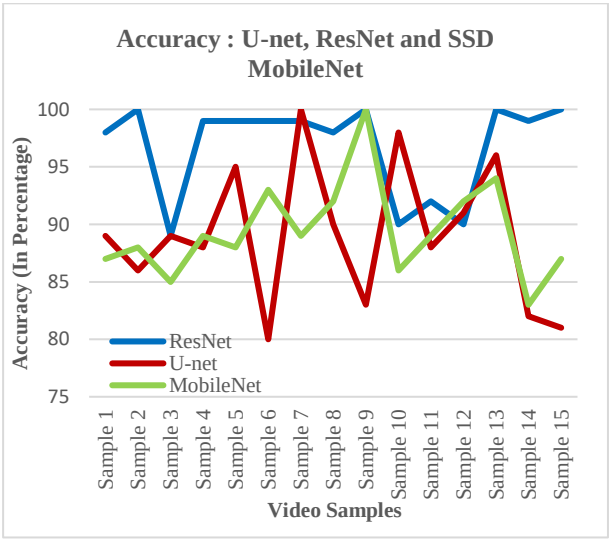

Fig. 9: The Graph Shows the Accuracy Comparison of the Two Networks U-Net Resnet and SSD Mobilenet Based on 13 Sample Footages Given as Input. the Accuracy Percentage of the Detections Performed for Each Footage by Both of the Networks Are Shown.

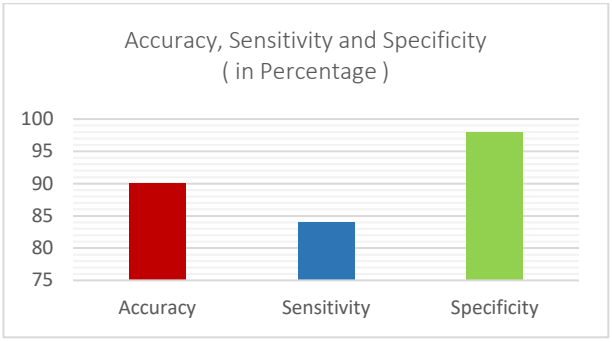

Fig. 10: Figure Shows the Accuracy, Sensitivity and Specificity Parameters for the Convolutional Neural Network Resnet, which Was the One That Shown Better Performance on the Comparisons Performed.

5. Conclusion

The proposed work is a GPU accelerated system for the fast, accurate and real-time detection of melanoma using deep learning. Three variants of deep convolution networks were used in the implementation of the system. Experimental results indicate that among the three systems implemented, the one with Residual Neural Networks are having better accuracy of 90.133% compared to that of system implemented with SSD MobileNet having an accuracy of 89.434% and U-net which is having an accuracy of 89.066%. Training phase of these system used images from International Skin Image Collaboration archives and testing were performed on human specimens. Detection of melanoma was performed by the system without any lesion segmentation or image pre-processing techniques. The system performed classification by processing frames which were extracted from the real-time input feed obtained by the system from the live camera. The system implemented performs real time detection from live input video feeds and is much better as compared to existing systems which uses image processing techniques to pre-process which takes melanoma images as input. So the proposed system performs video processing as compared to existing system which performs image processing. The GPU acceleration performed on the system increases the speed of computation and thus improves the ease of use and provides the capability of fast detection for the system. The results obtained from the system are comparable to the state-of-the-art systems giving an accuracy of 90.133%.

Acknowledgement

First and foremost, I sincerely thank the 'God Almighty' for blessing me with his grace. I wish to place on record my profuse sense of gratitude and sincere thanks to Dr. Surekha Mariam Varghese, Head Of the Department, Computer Science Engineering, for her guidance, constant supervision, encouragement and support throughout the period of this thesis work. I owe my sense of gratitude and sincere thanks to Prof. Elizabeth Isaac for her guidance, constant supervision, encouragement and support throughout the period of this thesis work. Finally, I would like to acknowledge the heartfelt efforts, comments, criticisms, co-operation and tremendous support given to me by my dear friends during the preparation of the project and also during the presentation without whose support this work would have been all the more difficult to accomplish.

References

[1] World Health Organization, Available at: http://www.who.int/uv/faq/skincancer/en/index1.html, accessed February 2018

[2] American Cancer Society, Available at: https://www.cancer.org/cancer/melanoma-skin-cancer/about/key-statistics.html, accessed February 2018

[3] A Ashfaq, I Marghoob and a Scope. The Complexity of Diagnosing Melanoma. In: Journal of Investigative Dermatology.2009, DOI:10.1038/jid.2008.388. https://doi.org/10.1038/jid.2008.388.

[4] NVIDIA Blog, Available at: https://blogs.nvidia.com/blog/2016/09/19/deep-learning-breast-cancerdiagnosis/

[5] S Affifi, H GholamHosseini and R Sinha. SVM classifier on chip for melanoma detection. In: Proceedings of the 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Seogwipo, South Korea 2017, DOI: 10.1109/EMBC.2017.8036814. https://doi.org/10.1109/EMBC.2017.8036814.

[6] S. Mustafa, A.B Dauda and M. Dauda. Image processing and SVM classification for melanoma detection. In: Proceedings of the International Conference on Computing Networking and Informatics (ICCNI), Lagos, Nigeria 2017, and https://doi.org/10.1109/ICCNI.2017.8123777.

[7] Z Waheed, a Waheed, M Zafar and F Riaz. An efficient machine learning approach for the detection of melanoma using dermoscopic images. In: Proceedings of the International Conference on Communication, Computing and Digital Systems (C-CODE), and Islamabad, Pakistan 2017, https://doi.org/10.1109/C-CODE.2017.7918949.

[8] An An Ali and H Al-Marzouqi. Melanoma Detection Using Regular Convolutional Neural Networks. In: Proceedings of the International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, and United Arab Emirates.2017, https://doi.org/10.1109/ICECTA.2017.8252041.

[9] M An Elahi, a Shahzad, M Glavin, E Jones and M O'Halloran. GPU Accelerated Confocal Microwave Imaging Algorithms for Breast Cancer Detection. In: Proceedings of ninth European Conference on Antennas and Propagation (EuCAP), Lisbon, Portugal.2015, p. 447-458.

[10] N R Abbasi, H M Shaw, D S Rigel, R J Friedman, W H McCarthy, I Osman, A W Kopf and D Polsky. Early Diagnosis of Cutaneous Melanoma Revisiting the ABCD Criteria, Journal of American Medical Association (JAMA), Vol. 292, No. 22, 2004, p: 2771-2776. https://doi.org/10.1001/jama.292.22.2771.

[11] O Ronneberger, P Fischer and T Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation, Journal of Computer Vision and Pattern Recognition, 2015, p: 1-8.

[12] K He, X Zhang, S Ren and J Sun. Deep Residual Learning for Image Recognition, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, United States of America.2016, DOI: 10.1109/CVPR.2016.90. https://doi.org/10.1109/CVPR.2016.90.

[13] E Smistad, T L Falch, M Bozorgi, A C Elster, F Lindseth. Medical Image Segmentation on GPUs - A Comprehensive Review,Journal of Medical Image Analysis, Elsevier, Vol. 20, 2015, p: 1-18. https://doi.org/10.1016/j.media.2014.10.012.

[14] N G Yadav. Detection of Lung Nodule using Content based Medical Image Retrieval, International Journal of Electrical, Electronics and Data Communication,2013, p: 2320-2084.

[15] M Birk, S Koehler, M Balzer, M Huebner, N V Ruiter and J Becker. FPGA based Embedded Signal Processing for 3D Ultrasound Computer Tomography, In Proceedings of 17th IEEE Real Time Conference (RT) 2011, p: 810-820.

[16] S Sarraf and G Tofighi. Deep Learning-based Pipeline to Recognize Alzheimer's disease using fMRI Data, In: Proceedings of IEEE Future Technologies Conference 2016, p: 816-820.

[17] K Pogorelov, M Riegler, P Halvorsen, P T Schmidt, C Griwodz, D Johansen, S L Eskeland and T de Lange. GPU-accelerated Real-time Gastrointestinal Diseases 67 Detection, Computer-Based Medical Systems (CBMS), In: Proceedings of IEEE 29th International Symposium, 2016, p: 251-265.

[18] Y Jia, E Shelhamer, J Donahue, S Karayev, J Long, R Girshick, S Guadarrama, and T Darrell. Caffe: Convolutional architecture for fast feature embedding, In: Proceedings of the ACM International Conference on Multimedia, 2014, p: 675 678. https://doi.org/10.1145/2647868.2654889.

[19] M Riegler, K Pogorelov, P Halvorsen, T de Lange, C Griwodz, P T Schmidt, S L Eskeland, and D Johansen. EIR - Efficient computer aided diagnosis framework for gastrointestinal endoscopies, In: Proc. of CBMI, 2016, p: 213-221.

[20] Y Wang, W Tavanapong, J Wong, J H Oh, and P C de Groen. Polypalert: Near real-time feedback during colonoscopy, Computer methods and programs in biomedicine, no. 3, 2015, p: 415-422.

[21] M Riegler, K Pogorelov, J Markussen, M Lux, H K Stensland, T de Lange, C Griwodz, P Halvorsen, D Johansen, P T Schmidt and S L Eskeland. Computer aided disease detection system for gastrointestinal examinations, In: Proc. of MMSys, 2016, p: 198-200.

[22] K Pogorelov, M Riegler, J Markussen, H Kvale Stensland, P Halvorsen, C Griwodz, S L Eskeland, and T de Lange. Efficient processing of videos in a multi-auditory environment using device lending of GPUs, In: Proc. of MMSys, 2016, p: 351-360. https://doi.org/10.1109/TITB.2012.2226595.
[23] Y Wang, W Tavanapong, J Wong, J Oh and P C deGroen. Near-realtime retroexion detection in colonoscopy, IEEE Journal of Biomedical and Health Informatics, vol. 17, no. 1, 2013, p: 143–152.
[24] R Nawarathna, J Oh, J Muthukudage, W Tavanapong, J Wong, P C De Groen, and S J Tan. Abnormal image detection in endoscopy videos using a lter bank and local binary patterns, 2014, p: 505-515.
[25] Y Wang, W Tavanapong, J Wong, J Oh, and P C de Groen. Computer-aided detection of retroexion in colonoscopy, In: Proc. of IEEE International Symposium on Computer-Based Medical Systems (CBMS), 2011, p: 1-6
[26] S Ioffe and C Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167, 2015.
[27] A G Howard, M Zhu, B Chen and D Kalenichenko. MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. Computer Vision and Pattern Recognition. ArXiv: 1704.04861v1, 2017.