# Exploring Deep Learning Networks for Tumour Segmentation in Infrared Images

Aman Dalmia[1,2], Siva Teja Kakileti[1], Geetha Manjunath[1]

[1]*Niramai Health Analytix, Koramangala, 560030, Bengaluru, India*
[2]*Indian Institute of Technology Guwahati, Amingaon, 781039, Guwahati, India*
*Contact Email: geetha@niramai.com*

## Abstract

Tumour segmentation is a crucial step in automated analysis of thermal images for early detection of breast cancer. Existing algorithms use clustering, thresholding and active contour techniques that rely on low-level image features generated from manual feature engineering. However, recent advances in deep learning, especially Convolutional Neural Networks (CNNs), are slowly making them the de-facto method for automated image analysis. In this paper, we explore various CNN architectures for semantic segmentation starting from naive patch based classifiers to more sophisticated ones including several variations of the encoder-decoder architecture for detecting the hotspots in the thermal image. We also show that encoder decoder architectures perform better when compared to patch based classifiers in terms of accuracy, dice index, jaccard index and inference time even with small thermal image datasets.

## 1. Introduction

Breast cancer is the most common type of cancer among women around the world. While mammography is widely used as a screening mechanism for breast cancer, it has various disadvantages with the major ones being that of lower sensitivity for younger women along with the risk of causing breast cancer because of repeated exposure to harmful radiation [23,24]. Nowadays, IR imaging has become an increasingly popular diagnostic tool to detect various diseases [16-18,28] and has been shown to reveal diagnostic and prognostic [29] information, lacking in other methods [25]. Due to the advent of high resolution thermal cameras [16-18], thermography is being re-considered as an alternate breast-imaging modality for early detection. In a study conducted by Gautherie et al.[1], It is found that an abnormal breast thermogram could be the earliest sign of tumor growth [1]. Due to the increased metabolism, a breast cancer tumor is at a higher temperature than the rest of the body. In [15], it is claimed that the use of abnormal hotspots for feature extraction could result in achieving high accuracy in differentiating malignancy and benign tumors when compared to whole breast features. In addition, localising the tumor spatially in the thermal image is needed for the diagnosis and prognosis as well. Hence, it is essential to segment the potential tumor regions for accurate classification of malignancy from thermograms.

Machine Learning is a class of artificial intelligence (AI) algorithms that uses mathematical techniques to model a decision problem that is usually performed by an expert. Traditional machine learning techniques involve a lot of manual feature-engineering, where the features, $f(X)$, has to be extracted from the data, $X$, with the help of domain and statistical knowledge. These extracted features are then fed along with the corresponding labels, Y, into the machine learning model for training as well as prediction. A major problem that has been with this approach is the requirement of expert domain knowledge to extract meaningful features from the dataset. To circumvent this problem a new class of representation learning algorithms was born. This class of algorithms called deep learning has become very popular in the past decade. Deep Learning models follow the general pattern of multiple levels of abstraction of input data formed from the combination of various non-linear modules to finally obtain learned features. Although these algorithms had been proposed long before, they have gained popularity only in recent years due to the availability of larger datasets and higher computation power in the form of powerful Graphic Processing Units (GPUs). Convolutional Neural Network (or CNN) is a type of deep learning model that observes an image in the form of patches and aggregates the inference made from each patch to make the final prediction. They are the current state-of-the-art for many computer vision problems like image recognition, classification, segmentation, localization, etc. But traditionally, a major problem with CNNs is the need for a large amount of training data to avoid the problem of overfitting and ensure generalization. Since getting vast amounts of data is especially hard in medical imaging, the

use of CNNs had been quite limited in this domain. Several efforts have been made to alleviate the data dependency of CNNs by proposing various regularization techniques to evade going into the overfitting regime. One such regularization technique called data augmentation, where the training set is artificially increased has proved to be of significant importance for tackling the overfitting problem in medical imaging applications.

In this paper, we explore and compare the use of sophisticated deep learning techniques for tumour segmentation in the context of breast thermograms. In section **2**, we discuss previous work done in tumor segmentation. Section **3** describes the various neural network architectures and the proposed modifications for their application to breast thermal images. Section **4** details our experimentation along with different data augmentation techniques needed for medical images when the amount of training data is less. We conclude in Section **5**, showcasing our results and comparing the various architectures described in Section **3**.

### 2. Related Work

Most of the prior work on breast cancer detection involves the use of textural features extracted from the entire image. The major drawback being localisation which is critical for the diagnosis of tumor. There are very few techniques in the literature that analyse the tumor region separately. Clustering techniques like fuzzy c-means and k-means have been proposed for segmenting the tumor region in breast thermal images [14]. In [26-27], authors employed different energy functions to detect the tumor region using active contours technique. Finally, [13,15] discusses the use of fusion of multiple adaptive thresholding techniques to identify the warm spots and hot spots in a thermal image where temperature thresholds are decided by certain parameters that are chosen to maximize the linear combination of sensitivity and specificity. The inherent disadvantage of all these techniques is their strong dependency on the thresholds and their consideration of low-level features for segmentation. Also, unlike normal images where the values are quantized between 0 and 255, thermal images of the breast range anywhere from 15°C to 40°C. Use of fixed thresholds for segmentation in those conditions lead to poor performance owing to varying range of body temperatures and requires manual intervention when the temperature ranges in a new dataset are different from the original training set.

### 3. Convolutional Neural Networks for Tumour Localization

We propose the use of CNNs to learn meaningful semantic representation of hotspots and explore the use of multi-layered CNN for detection of hotspots in infrared breast thermal images. Each layer in a CNN can be individually approximated to an edge based filter or a complex threshold filter and thus having more layers can help in abstraction of higher level features for better segmentation. We discuss the functionalities of these architectures and how they learn the properties that are relevant to abnormal hotspot alone by ignoring pixels corresponding to spurious hotspots that are usually generated from thresholding technique. As a ground truth for comparison, we consider hotspot segmentation based on adaptive fusion thresholding [13,15] that is known to detect hot spots in thermal images with high accuracy, but requires manual intervention in deciding temperature range. We use these hotspots generated after manual intervention to fit a CNN that could learn these features without the need of manual intervention. In the recent years, there are several CNN architectures that have been proposed for medical image segmentation. These architectures can be broadly categorized into a) Point-wise classification and (b) End-to-End CNN. We discuss each of these class of architectures for breast tumour localization in detail below.

#### 3.1 Point-wise classification:

The first approach to semantic segmentation in medical images is the classification of each pixel [11] in the image by running a sliding window across the whole image and taking a patch around the pixel under consideration with the output being a single value denoting the class to which the pixel belongs. We consider two separate architectures under this branch of architectures: VGGNet and InputCascadeCNN. The loss function for both the architectures is the binary cross-entropy.

### 3.1.1 VGGNet

VGGNet [19] has won the ImageNet localisation challenge in 2014 for detecting and localising different objects in the image with high accuracy using a larger network consisting of 19 layers. The network takes a patch as input, and outputs the classification result for each patch. The main innovation of this architecture is its simplicity and depth, where 3x3 convolutions are repeatedly applied in contrary to the preceding architectures like AlexNet, where filters of different sizes like 7x7, 11x11 were used. VGGNet showed the effectiveness of small filters owing to the fact that two consecutive 3x3 convolutions have an effective receptive field of 5x5. Also, the number of filters double after each max pooling operation to transfer the information learnt from previous layers..

### 3.1.2 InputCascadeCNN

This architecture proposed in [2] is a multi-channel cascade CNN applied for brain tumor segmentation. The motivation behind this architecture is to have two different pathways: a local pathway, which looks at a small region around the pixel, and a global pathway, which looks at a large portion of the image, and combination of their inferences can be used for making the final prediction. The authors point out a major problem with widely used CNN architectures is that the prediction of a pixel in consideration can't be influenced by the predictions of its surrounding pixels. They get rid of this problem by using a cascaded architecture, where two patches of different sizes centered around a pixel in consideration are passed as inputs to the network. The global patch passes through the network first, whose output is concatenated with the local smaller patch and then used as the input to the same network. Maxout [20] is used as the activation function between the convolutional layers is instead of ReLU owing to the fact that it leads to a more efficient model averaging when used with dropout. This model is named *InputCascadeCNN* as the output from the first stream is concatenated to the smaller input patch of the second stream. Fig .1 shows the modified architecture for hotspot segmentation in breast thermal images.
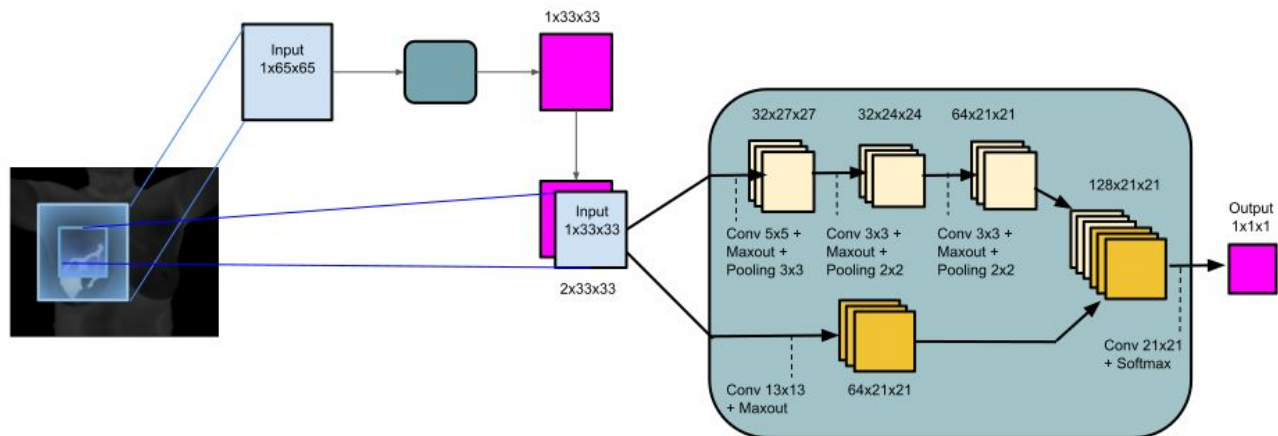


***Fig. 1***: *Modified  InputCascadeCNN[2] for hotspot segmentation in breast thermal images*

### 3.2 End-to-end CNN

There are various problems associated with the approaches mentioned above. A major one being that it involves a lot of redundant computations because of overlapping windows. In the end-to-end CNN approach, the input is the entire image in consideration, and the output would be the segmentation mask itself. A first approach to this form of segmentation was proposed in [3], which introduces a convolutional implementation of the sliding window approach. However, variants of fully convolutional networks (FCN) involving an encoder-decoder architecture as proposed in [10] are giving the state-of-the-art results in semantic segmentation today. We describe two such variants of the FCN which have been recently proposed: U-Net [21] & V-Net [4]. Unlike conventional losses used for image classification, they proposed the dice coefficient as the loss function to compare the similarity between actual and predicted outputs during training.

### 3.2.1 U-Net

The U-Net [21] architecture builds on the encoder-decoder architecture where the encoder transforms the image to a lower dimension space by extracting high-level features down the network and the decoder network then reconstructs the output of the desired shape. The encoder network consists of multiple sub-stages with each stage consisting of two 3x3 convolutional layers followed by a Max-Pooling layer to perform the downsampling operation. The number of filters are doubled in each subsequent stage to capture wider set of features. The decoder network is similar to the encoder network, with the downsampling operation being replaced by a transpose convolution, also called strided convolution or deconvolution, operation. The corresponding output of the encoder network at each stage is also concatenated to the input of each stage in the decoder network as shown in Fig. 2. This architecture relies heavily on data augmentation for its performance, which is explained in Section **4.2**.
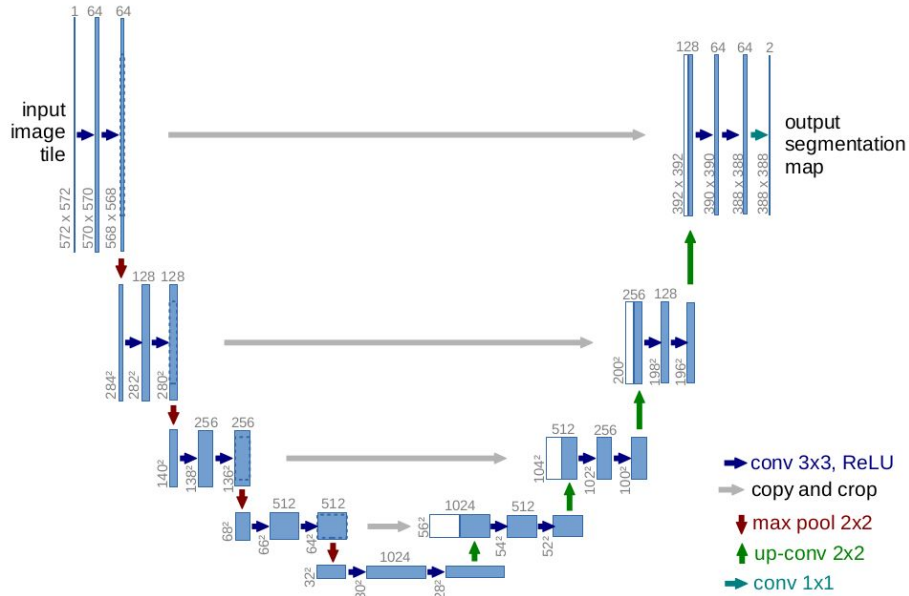


**Fig. 2**: *U-net architecture proposed by Ronneberger et.al in [21] showing different stages of encoder and decoder*

### 3.2.2 V-Net

The V-Net [4] architecture as shown in Fig .3 is also an encoder-decoder architecture inspired from [10], but it uses residual connections in all its stages to speed up the training procedure [22], along with the freedom of using a deeper network. Also, the architecture uses strided convolutions for downsampling, instead of Max-Pooling, which has drawn heavy criticism from the deep learning community due to its lack of interpretability. PReLU is used as the non-linearity instead of Rectified Linear Unit (ReLU). A drawback in the U-Net architecture was that the positive and negative regions were weighted equally. This is generally not the case especially in medical datasets where there will high imbalance between positive and negative samples. This results in the model being biased towards the negative pixels. Instead of resorting to optimizing the dice coefficient coupled with sample re-weighing, the authors proposed to optimize the derivative of the dice coefficient with respect to each voxel $p_j$ as shown in the below equation, where $p_i \in P$, the predicted segmentation map and $g_i \in G$, the ground truth mask.

$$\frac{\partial D}{\partial p_j} = 2 \left[ \frac{g_j \left( \Sigma_i^N p_i^2 + \Sigma_i^N g_i^2 \right) - 2 p_j \left( \Sigma_i^N p_i g_i \right)}{\left( \Sigma_i^N p_i^2 + \Sigma_i^N g_i^2 \right)^2} \right]$$

## *4. Experimentation*

Our dataset consists anonymized data of 180 subjects, with 5 images per subjects corresponding to frontal (0°), oblique (± 45°) and lateral (± 90°) views. Among the 180 subjects, 120 subjects are captured using FLIR T650Sc camera, which has a thermal sensitivity of 0.02°C and spatial resolution of 480 x 640. The remaining images are captured using A315 camera with a spatial resolution of 240 x 320 and a thermal sensitivity of 0.05°C. Each thermal image represents raw temperature data emitted from the body. Out of 180 subjects, we randomly choose 150 subjects for training and 30 subjects for validation. Thus, the validation set resembles the real-world task closely as the model doesn't see any image of the validation subjects during training.
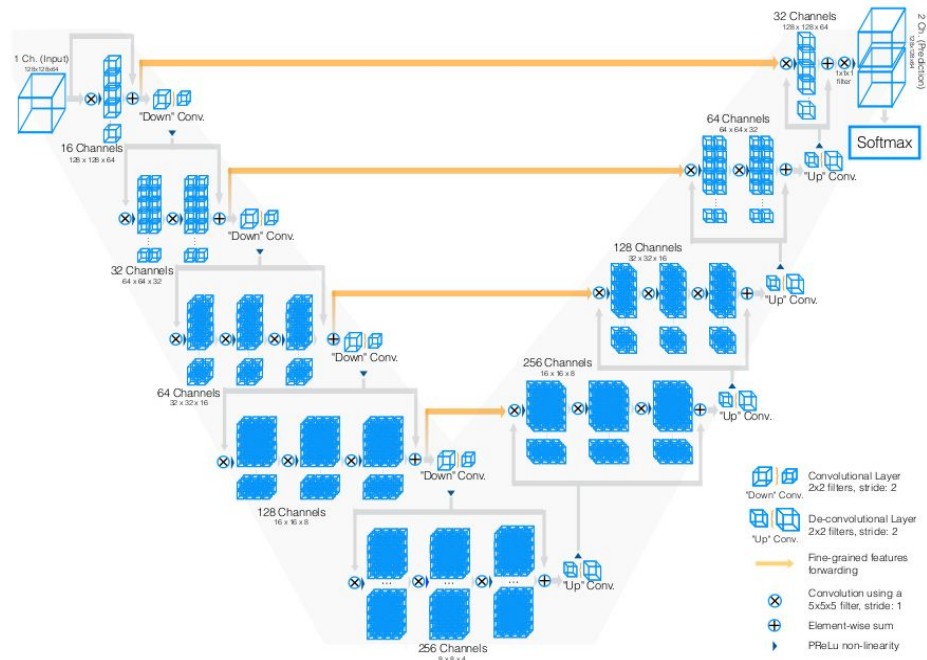


***Fig. 3****: V-net architecture proposed by* Milletari *in [4] with identity mapping in each stage  of encoder and decoder*

For the point-wise classification model, upon manually observing many subjects, we generally found the size of the largest hot spot as 120x120 patch for a 480x640 image and  60x60 patch for a 240x320 image. Owing to the computational cost of taking a patch across each of the pixels in these high resolution images, we down-scaled the images to 120x160 resolution and the patch size is correspondingly rescaled to 33x33 with 3 pixels representing the margin of error in x & y direction. Now, since hot spots occupy a relatively small portion of the image, we have high imbalance between negative samples than positive samples. From this imbalance pool of positive and negative samples, we create batches with equal number of negative and positive samples. This is done to make sure that the network receives a balanced dataset so that binary cross-entropy becomes a valid loss function.

On other hand, End-to-End segmentation models require the whole image as input. As the dataset is too small to learn, it could result in overfitting. Hence, data augmentation plays a very important role in generalizing these models to unseen subjects. This is mentioned in more detail in below sub sections along with other details of preprocessing, training and implementation.

### 4.1 Preprocessing:
Thermal images are unlike usual medical images where the images are quantized between gray levels. Since we are capturing raw temperature data, the values can even change with the time of capture, emotional state of subject and even depend on the surrounding temperature. Hence, appropriate normalization needs to be applied for the model to be invariant

to such changes. Firstly, foreground segmentation is applied to detect body pixels using Otsu's [7] threshold. To normalise the image, we subtract the image with the minimum body temperature and divide it by the body temperature range, the difference between maximum body temperature and minimum body temperature.

### 4.2 Data augmentation:

A major bottleneck for the application of deep learning to medical imaging is the lack of enough data. In our case too, we only have 150 subjects to train with each of them having 5 images corresponding to left lateral, right lateral, frontal, left oblique and right oblique. However, this doesn't mean that we have 750 images for all architectures, as we only count the images that actually have a hotspot for pointwise architectures and are finally left with approximately 450 images. Clearly, both 750 and 450 images are quite small for training a CNN. Hence, data augmentation plays a very important role for us. The model should learn to be invariant to affine transformations like rotation, horizontal flipping, horizontal and vertical shifts, shear and zoom. Apart from these, random elastic transformations seem to be of great importance for biomedical imaging use cases. We generate smooth deformations as in [21] on a 3x3 grid using a Gaussian distribution of 10 pixels standard deviation. These transformations artificially increases the dataset and it has been shown to also act as a regularizer, both of which are desirable properties and hence, it is extremely important for this task.

### 4.3 Training:

Batch Normalization [5] help in speeding up the training process, providing the freedom to employ a larger learning rate and also acting as a regularizer. Batch Normalization was applied after the convolutional layers, preceding the activation layer, as suggested in the original paper [5]. Apart from this implicit regularization, various explicit regularizers like L1 and L2-norms were utilized at each layer to prevent overfitting in VGGNet, Max-out[20] was used as regularizer in InputCascadeCNN and Dropout [15] was applied as the regularizer for the End-to-End CNNs. We used the Adam [6] optimizer for all our architectures. A learning rate of 0.0001 was used for the point-wise classification models and 0.00001 for the end-to-end segmentation models. We observed that training the model from scratch performed better than using the pre-trained weights.

### 4.4 Implementation:

All the models were implemented using the Keras [8] framework, with TensorFlow as the backend. For performing hyperparameter tuning, we used the Tree of Parzen Estimators [9] technique, a variation of Bayesian Optimization. The framework, Hyperas, which is wrapper for Keras, was used for this task. All the simulations are run on the NVIDIA GeForce GTX TITAN X (12GB) GPU.

### *5. Results*

We present our results in two-fold: First, we compare the computational feasibility of the different architectures followed by their performance on our breast thermal dataset. Table 1 compares the number of parameters and inference time of the mentioned architectures. It can be seen that the encoder-decoder based Fully Convolutional Network architectures take the least time for inference on one image, even though they have higher number of parameters. Inference time mentioned in the table indicates the time taken by a model to run a feed-forward once on its corresponding input image. While the inputs for End-to-End CNNs are the entire image, the point-based models take patches around the pixel as the input. So, if we intend to find the prediction map for the entire image, we'd have to run the point-based models multiple times to cover all the patches. Hence, the total prediction time is much larger compared to end-to-end FCN models. In addition, for point-wise classifications, the output is a single scalar value that indicates whether the center pixel of the patch belongs to a hot-spot or not. On other hand, for U-Net and V-Net, the output is the final segmentation map. Thus, accuracy is used as the metric for point-wise architectures, whereas dice-coefficient is preferred for encoder decoder architectures. The corresponding training and validation curves can be seen in Fig. 4.

**Table 1.** Comparison of the number of parameters and the training times (using 20% of GPU) of the different architectures

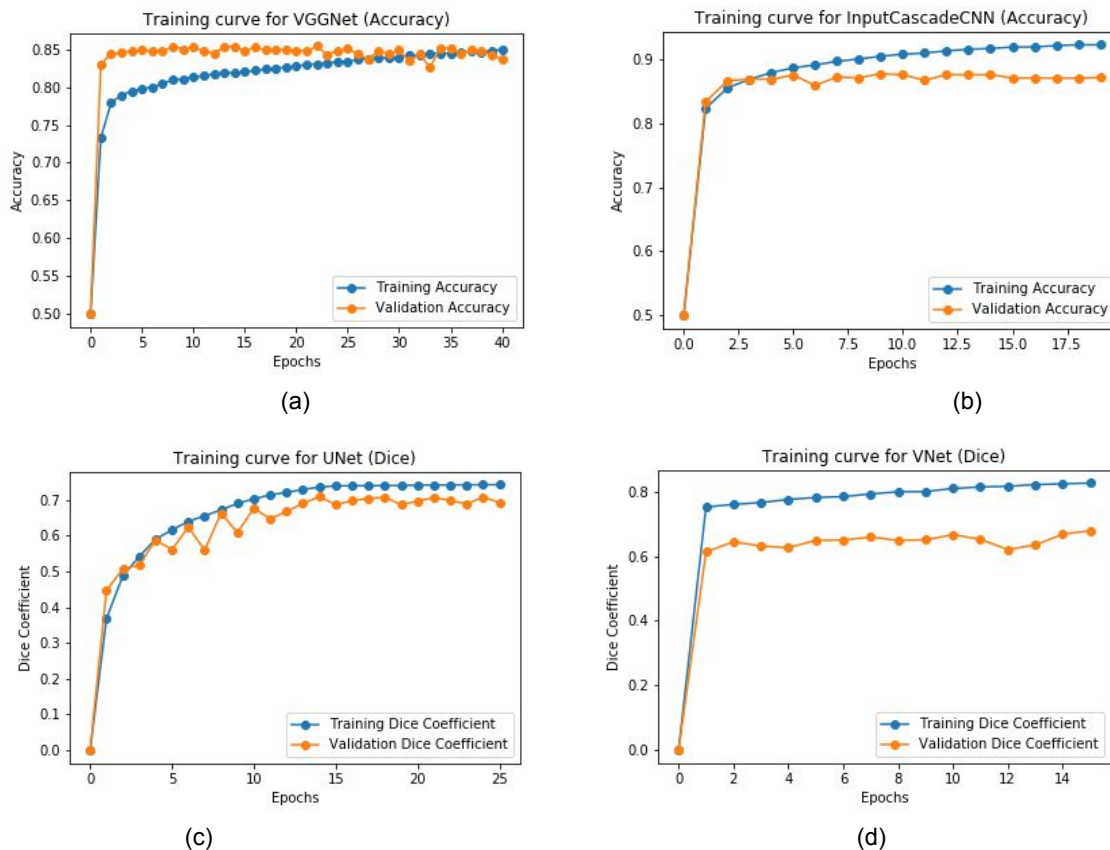| Architecture | Number of Parameters | Inference time per image (in seconds) |
|---|---|---|
| VGGNet | 50,625 | 0.6 |
| InputCascadeCNN | 112,898 | 95 |
| U-Net | 485,009 | **0.2** |
| V-Net | 6,301,697 | 0.2 |



(a)  (b)

(c)  (d)

*Fig 4. Training curves showing the changes in accuracy over epochs for (a) VGGNet and (b) InputCascadeCNN & the changes in dice-coefficient over epochs for (c) U-Net & (d) V-Net*

Table 2 compares the various architectures on 3 performance metrics: accuracy, dice-coefficient and jaccard Index. We found that though VGGNet was able to fit the training set perfectly (overfitted), it was able to produce validation accuracy close to 0.7 without applying any kind of regularization. However, upon applying the regularization techniques mentioned above, the maximum validation accuracy has increased from 0.7 to 0.85. The best performing architecture for the multi-channel cascaded CNN was a variant of the original architecture as described in Fig.1, where we used 3 convolutional layers instead of 2, with filter sizes 5x5, 3x3 and 3x3 respectively. We obtained a training accuracy of 0.92 corresponding to the best validation accuracy of 0.87 and a test accuracy of 0.86. However, these results were obtained when an equal number of true and false patches as mentioned. For both the end-to-end segmentation models, the training dice coefficient plateaus at about 0.8 with their best validation dice coefficient close to 0.7.
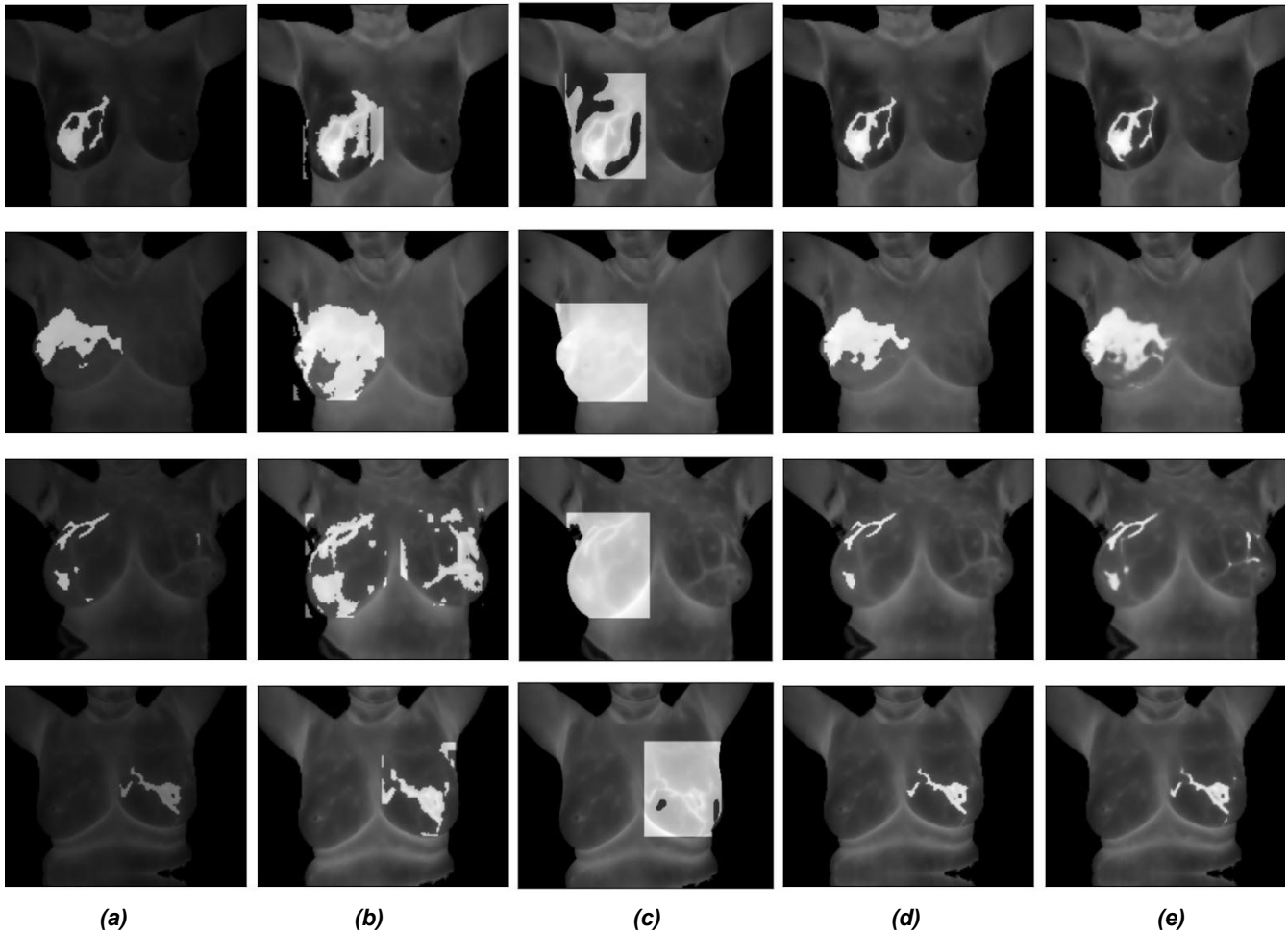
|     |     |     |     |     |
| --- | --- | --- | --- | --- |
| *(a)* | *(b)* | *(c)* | *(d)* | *(e)* |

***Fig 5.**. Comparison between the actual and predicted segmentation maps.* (a) Ground truth *(b) VGGNet (c) InputCascadeCNN (d) U-Net and (e) V-Net for 4 different subjects. It shows that the problem faced by the FCN models with the encoder-decoder architecture is that of false negatives whereas the one faced by point-based models is of false positives. The examples also make it clear that the problem of false positives in the point-based models is much more prominent compared to the problem of false negatives in the FCN models.*

Based on the above mentioned metrics, it can be clearly seen that encode-decoder based architectures are more suited for the semantic segmentation of hot spots in thermal images compared to patch based classifications. To get a more intuitive understanding of the predictions made by the various models described above, we superimpose the predicted masks on the original image and compared them with the ground truth labels for each of the four discussed models as shown in Fig. 5. From the heatmaps for each subject, it can be clearly seen that the segmentation masks predicted by VGGNet is reasonable close to the ground truth annotations whereas that of InputCascadeCNN fails miserably classifying almost the entire breast region as a hot-spot. Typically, CNNs suffer from the problem of false negatives where it misses out some portions of the hot spot region. But the errors in the segmentation map produced by the point-based models are attributed to false positives. On the contrary, segmentation maps produced by the End-to-End CNNs suffer from few false negatives compared to very few false positives. But the overall error in the segmentation maps produced by the point-based models is much larger than End -to-End CNNs mainly due to the large number of false positives. Based on the above results coupled with the fact that

Encoder-Decoder architectures are the state-of-the-art in computer vision for semantic segmentation tasks, we infer that End-to-End CNN architectures are more suited to the hot spot segmentation task than point-wise classification models.

***Table 2.*** Comparison of Accuracy, Dice and Jaccard indices for different architectures. For the point-based models, the accuracy values are reported using a balanced dataset where each batch consists of equal number of negative and positive samples.

| Architectures Metrics | VGGNet | InputCascadeCNN | U-Net | V-Net |
|---|---|---|---|---|
| Accuracy (Train, Validation) | 80.1, 81.6 | 92.1, 87.6 | 99.5, 99.5 | **99.6**, **99.6** |
| Dice-coefficient (Train, Validation) | 34.6, 45.4 | 31.8, 30.2 | 76, 75.8 | **79.6**, **79.9** |
| Jaccard Index (Train, Validation) | 20.9, 29.5 | 15.4, 16.3 | 61.3, 60.9 | **66.2, 66.6** |

### *6. Conclusion*

The segmentation of hot spots in a thermal image is a tough problem mostly due to the un-availability of large breast thermal datasets, lack of normalised data and the dependency of captured thermal images on ambience, emotional and physical state of the subject. In this paper, we have explored the performance of various deep learning networks that produced a good training performance in terms of accuracy, dice coefficient and jaccard index along with a decent validation performance even on a small breast thermal dataset. We believe that larger dataset in combination with parameter tuning could help the model to efficiently generalize to unseen examples and thereby, producing higher training and validation results. We also showed the importance of loss function in improving the results especially in the case of V-net architecture, where it alleviated the need to assign different weights to the positive and negative classes. Additionally, we believe these techniques could also be used for other infrared imaging use cases like detecting the defects in solar panel, identifying high consumption regions in a chipset etc.

### REFERENCES

[1] M. Gautherie, "Thermobiological assessment of benign and malignant breast diseases," Am. J. Obstet. Gynecol., vol. 8, no. 147, pp. 861-869, 1983.

[2]Havaei, Mohammad, Axel Davy, David Warde-Farley, Antoine Biard, Aaron Courville, Yoshua Bengio, Chris Pal, Pierre-Marc Jodoin, and Hugo Larochelle. "Brain tumor segmentation with deep neural networks." Medical image analysis 35 (2017): 18-31.

[3] Sermanet, Pierre, David Eigen, Xiang Zhang, Michaël Mathieu, Rob Fergus, and Yann LeCun. "Overfeat: Integrated recognition, localization and detection using convolutional networks." arXiv preprint arXiv:1312.6229 (2013).

[4] Milletari, Fausto, Nassir Navab, and Seyed-Ahmad Ahmadi. "V-net: Fully convolutional neural networks for volumetric medical image segmentation." In 3D Vision (3DV), 2016 Fourth International Conference on, pp. 565-571. IEEE, 2016.

[5] Ioffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by reducing internal covariate shift." arXiv preprint arXiv:1502.03167 (2015).

[6] Kinga, D., and J. Ba Adam. "A method for stochastic optimization." In International Conference on Learning Representations (ICLR). 2015.

[7] Otsu. N. "A threshold selection method from gray-level histograms." IEEE transactions on systems, man, and cybernetics 9, no. 1 (1979): 62-66.

[8] Chollet, F. (2017). "Keras (2015)". URL http://keras.io

[9]Bergstra, James S., Rémi Bardenet, Yoshua Bengio, and Balázs Kégl. "Algorithms for hyper-parameter optimization." In Advances in neural information processing systems, pp. 2546-2554. 2011.

[10]Long, Jonathan, Evan Shelhamer, and Trevor Darrell. "Fully convolutional networks for semantic segmentation." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 3431-3440. 2015.

[11] Salas, Yainuvis Socarrás, David Vázquez Bermudez, Antonio M. López Peña, David Gerónimo Gomez, and Theo Gevers. "Improving hog with image segmentation: Application to human detection." In International Conference on Advanced Concepts for Intelligent Vision Systems, pp. 178-189. Springer, Berlin, Heidelberg, 2012.

[12] Srivastava, Nitish, Geoffrey Hinton, Alex Krizhevsky, Ilya Sutskever, and Ruslan Salakhutdinov. "Dropout: A simple way to prevent neural networks from overfitting." The Journal of Machine Learning Research 15, no. 1 (2014): 1929-1958.

[13] Venkataramani, Krithika, Lalit K. Mestha, L. Ramachandra, S. S. Prasad, Vijay Kumar, and Priyanka J. Raja. "Semi-automated breast cancer tumor detection with thermographic video imaging." In Engineering in Medicine and Biology Society (EMBC), 2015 37th Annual International Conference of the IEEE, pp. 2022-2025. IEEE, 2015.

[14] EtehadTavakol, M., et al "Application of K-and fuzzy c-means for color segmentation of thermal infrared breast images." *Journal of medical systems* 34.1 (2010): 35-42.

[15] Madhu, Himanshu, et al. "Extraction of medically interpretable features for classification of malignancy in breast thermography." *Engineering in Medicine and Biology Society (EMBC), 2016*

[16] D. A. Kennedy, T. Lee and D. Seely, "A comparative review of thermography as a breast cancer screening technique," Integrative Cancer Therapies, vol. 8, no. 1, pp. 9-16, 2009.

[17] J. R. Keyserlingk, P. D. Ahlgren, E. Yu, N. Belliveau and M. Yassa, "Functional infrared imaging of the breast," IEEE Engineering in Medicine and Biology Magazine, vol. 19, no. 3, pp. 30-41, 2000.

[18] Kakileti, Siva Teja, Geetha Manjunath, Himanshu Madhu, and Hadonahalli Venkataramanappa Ramprakash. "Advances in Breast Thermography." In New Perspectives in Breast Imaging. InTech, 2017.

[19] Simonya.K , and Andrew .Z. "Very deep convolutional networks for large-scale image recognition." arXiv preprint:1409.1556 (2014).

[20] I. J. Goodfellow, D. Warde-Farley, M. Mirza, A. Courville, and Y. Bengio. "Maxout networks." In Proceedings of the 30th International Conference on Machine Learning, pages 1319– 1327. ACM, 2013.

[21] Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 234-241). Springer, Cham.

[22] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[23] Yaffe, Martin J., and James G. Mainprize. "Risk of radiation-induced breast cancer from mammographic screening." Radiology 258, no. 1 (2011): 98-105.

[24] de González, A. Berrington, and G. Reeves. "Mammographic screening before age 50 years in the UK: comparison of the radiation risks with the mortality benefits." British journal of cancer 93, no. 5 (2005): 590.

[25] Y. Poussart, M. Frize, R. Roberge, "The re-evaluation of thermography in breast cancer detection by new image enhancement techniques", Terry Fox Foundation Fredericton NB Canada Research Report, 1988

[26] Golestani, N., M. EtehadTavakol, and E. Y. K. Ng. "Level set method for segmentation of infrared breast thermograms." EXCLI journal 13 (2014): 241.

[27] Zadeh, Hossein Ghayoumi, Javad Haddadnia, Omid Rahmani Seryasat, and Sayed Mohammad Mostafavi Isfahani. "Segmenting breast cancerous regions in thermal images using fuzzy active contours." EXCLI journal 15 (2016): 532.

[28] Ng, E. Y. K., Chen, Y., and Ung, L. N., Computerized breast thermography: Study of Image segmentation and temperature cyclic variations. Intl. J. Med. Eng. Technol. 25:12–16, 2001.

[29] Kakileti, Siva Teja, Krithika Venkataramani, and Himanshu J. Madhu. "Automatic determination of hormone receptor status in breast cancer using thermography." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 636-643. Springer, Cham, 2016.