

Article

Deep Vision for Breast Cancer Classification and Segmentation

Lawrence Fulton ^{1,*}, Alex McLeod ², Diane Dolezel ³, Nathaniel Bastian ⁴ and Christopher P. Fulton ⁵¹ College of Health Professions, Texas State University, San Marcos, TX 78666, USA² McCoy College of Business, Texas State University, San Marcos, TX 78666, USA; am@txstate.edu³ Health Information Management, College of Health Professions, Texas State University, Round Rock, TX 78665, USA; dd30@txstate.edu⁴ United States Military Academy, West Point, NY 10996, USA; nathaniel.bastian@westpoint.edu⁵ United States Air Force, Palmdale, CA 93551, USA; christopher.fulton.6@us.af.mil

* Correspondence: lf25@txstate.edu; Tel.: +1-210-837-9977

Simple Summary: Breast cancer misdiagnoses increase individual and system stressors as well as costs and result in increased morbidity and mortality. Digital mammography studies are typically about 80% sensitive and 90% specific. Improvement in classification of breast cancer imagery is possible using deep vision methods, and these methods may be further used to identify autonomously regions of interest most closely associated with anomalies to support clinician analysis. This research explores deep vision techniques for improving mammography classification and for identifying associated regions of interest. The findings from this research contribute to the future of automated assistive diagnoses of breast cancer and the isolation of regions of interest.

Abstract: (1) Background: Female breast cancer diagnoses odds have increased from 11:1 in 1975 to 8:1 today. Mammography false positive rates (FPR) are associated with overdiagnoses and overtreatment, while false negative rates (FNR) increase morbidity and mortality. (2) Methods: Deep vision supervised learning classifies 299×299 pixel de-noised mammography images as negative or non-negative using models built on 55,890 pre-processed training images and applied to 15,364 unseen test images. A small image representation from the fitted training model is returned to evaluate the portion of the loss function gradient with respect to the image that maximizes the classification probability. This gradient is then re-mapped back to the original images, highlighting the areas of the original image that are most influential for classification (perhaps masses or boundary areas). (3) Results: initial classification results were 97% accurate, 99% specific, and 83% sensitive. Gradient techniques for unsupervised region of interest mapping identified areas most associated with the classification results clearly on positive mammograms and might be used to support clinician analysis. (4) Conclusions: deep vision techniques hold promise for addressing the overdiagnoses and treatment, underdiagnoses, and automated region of interest identification on mammography.

Keywords: deep vision; breast cancer; machine learning; region of interest detection



Citation: Fulton, L.; McLeod, A.; Dolezel, D.; Bastian, N.; Fulton, C.P. Deep Vision for Breast Cancer Classification and Segmentation. *Cancers* **2021**, *13*, 5384. <https://doi.org/10.3390/cancers13215384>

Academic Editors: Yi Zhao, Jinpu Yu, Haifeng Zhang, Min Hu and Jincheng Guo

Received: 13 October 2021

Accepted: 24 October 2021

Published: 27 October 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

An estimated 2.3 million women were diagnosed with breast cancer globally in 2020, and female breast cancer has surpassed lung cancer as the most diagnosed cancer in the world [1]. The odds of a female being diagnosed with breast cancer have increased from 11:1 in 1975 to 8:1 today [2]. Globally, breast cancer is the most prevalent type of cancer with the most disability-adjusted life years [3]. Age-adjusted rates show growth in breast cancer diagnoses of 0.3% per year [4,5].

In the United States, about 13% of women will be diagnosed with breast cancer over their lifetimes [6]. Although breast cancer fatality rates have declined 1% since 2013, likely due to advancements in treatment, it is still the second most fatal cancer diagnosis [6]. Fatality rates from breast cancer decreased 7% from 2002 to 2003 (possibly due to an

evidence-based reduction in the use of hormone replacement therapy [6,7]). However, the increased use of breast implants (48% increase from 2000 to 2018, the pre-COVID-19 era [8]) is linked to an increased risk of breast implant associated anaplastic large cell lymphomas (BIA-ALCL). These lymphomas are rare and difficult to detect [9], facts associated with a reduced probability of 5-year survival. The Scientific Committee on Health, Environment, and Emerging risks considers there to be moderate evidence of a causal link between textured breast implants and BIA-ALCL [10].

Mammography reduces breast cancer mortality and is thus an important diagnostic tool [11–13]. The implementation of breast cancer screening has resulted in the increase in ductal carcinoma in situ, a pre-invasive form of cancer that may or may not progress [14]. Misdiagnoses of mammography results in increased morbidity, mortality, stress, and costs. False positive rates (FPR), estimated in one study to be 121.2 per 1000 or 12.12% [15], are associated with overdiagnoses and overtreatment and must be addressed while simultaneously improving false negative rates (FNR) to reduce morbidity and mortality. One study estimates that sensitivity and specificity of digital mammography studies is 80% and 90%, respectively [16].

The use of human–computer interaction in mammography studies is becoming increasingly important [17]. Advances in machine learning (ML) have improved both positive predictive validity (PPV) and negative predictive validity (NPV) in breast cancer identification. Application of fast opposite learning weights, an ML technique that may improve classification performance, resulted in FNR of 9.9% and FPR of 11.94% in one study [18]. A study by Ertosun and Rubin used Convolutional Neural Network (ConvNet) achieved 0.78 precision in mammography classification [19]. Muramatsu et al. achieved 84% accuracy using ConvNet. ML techniques hold promise for improving PPV and NPV [20]. Salama and Aly were able to achieve impressive classification metrics (98% accuracy, 98% specificity, and 98% sensitivity) on small samples with augmented imagery [21].

Image segmentation attempts to divide an image into non-overlapping areas. Techniques that have been applied to this task include thresholding, watershed-based methods, graph-based methods, clustering, and region-based approaches [22]. Thresholding methods are often exceedingly basic, seeking to enhance grayscale images based on intensity values [23]. Smooth boundaries and unimodal histograms (pixel value) pose problems for this method [22]. Watershed-based methods consider the image is a contour map and seek to find the lowest/highest points [24] but are sensitive to noise and over-segmentation and are computationally expensive as gradient calculation is required [22]. Graph-based methods subset images using nodes and edges where the edges never overlap and have been used for region of interest (ROI) identification in breast cancer [25]. They are, again, computationally problematic for larger image sizes [26]. Clustering methods (K-means, hierarchical, etc.) seek to group like pixels together in order. These algorithms are sensitive to outliers, noises, and initial values [25]. Region-based methods seek to divide an image into homogenous regions but are sensitive to the initial region selected as the ‘seed’ [22].

Recent efforts have sought to use the output of a ConvNet classification model to identify ROIs on photographs without knowing where the actual lesion may exist. Techniques such as ‘guided backpropagation’ [27] and ‘deconv’ [28] have been used to identify image areas that activate the neurons of ML models. The gradient of the loss function with respect to an image representation (convolutional layer) can be used to pinpoint tissue segments that might be of concern even if the initial read of that study was negative without previously classified imagery (unsupervised learning). Recent efforts have shown that such methods are highly capable of producing bounding boxes for imagery [29]. It is therefore possible to identify regions of interest (ROIs) in medical imagery by simply ‘mining’ the layers of a ConvNet classification algorithm without additional segmentation analysis.

This study addresses the problems of overdiagnoses and overtreatment, as well as underdiagnoses with its associated increases in morbidity and mortality by (1) improving classification of mammography using supervised learning (learning from images where

the classification label is in the data) and (2) implementing guided backpropagation to paint gradient contours of the ROIs using unsupervised learning. This research supports efforts to reduce morbidity and mortality while simultaneously addressing overdiagnosis and overtreatment.

2. Materials and Methods

2.1. Data, Software, and Hardware

Data are publicly available from the Digital Database for Screening Mammography (DDSM) [30] and the Curated Breast Cancer Imaging Subset of DDSM [31] and provided by Google's Kaggle.com, accessed on 5 January 2021 [32]. This analysis is a precursor to follow-on work supported by a National Cancer Institute Data Transfer Use Agreement (PLCOI-742). The data consisted of 71,249 images and labels, 55,885 pre-designated for training and 15,364 reserved for testing. The image data were sized 299×299 (single channel grayscale) but were augmented to three-color ('RGB') for use in models by replicating the channel. The label data included dichotomous classification (0 for true negative, 1 otherwise). All analyses were performed in Anaconda Python 3.7 and are available on Github [33]. An in-kind high performance computing grant from Advanced Micro Devices (250 teraflops computing power) provided the computational power for model training. The observations (pixels) used for the image data only were $299 \text{ pixels} \times 299 \text{ pixels} \times 3 \text{ channels} \times 71,249 \text{ images} = 19.1 \text{ billion pixels}$.

2.2. Training, Validation, and Test Sets

The training set was further randomly subdivided into training and validation sets (80% and 20%, respectively) resulting in the final data sets of size 44,708 (training), 11,177 (validation), and 15,364 (test). By further splitting the training data, machine learning algorithms use the retained training data to estimate model performance on the validation set prior to estimation of the test set as a mechanism for preventing overfitting and to compare model hyperparameters. The final tuned model used both the training and validation sets to predict the pristine test set.

2.3. Image and Label Preprocessing

Each of the image intensity values were scaled between 0 and 1 (min-max scaling). Test data remained untransformed other than pixel scaling. Label data were complete.

2.4. Architecture

ConvNet models were built on the training images and performance calculated using the validation set. Then the models were used to classify the unseen but labeled test set. ConvNet is an ML network architecture capable of taking radiological images and making outcome predictions. ConvNets are often used for image denoising [34], as they preserve the important spatial relationships and features. Figure 1 is a basic ConvNet architecture.

2.5. Deep Vision Basics

An image is represented as an x-y mapping of pixel color intensity, z. For color images, there are three matrices that form an array, one each for red, blue, and green intensities. One could consider ConvNet to be a dimension reduction protocol that reduces the image size while retaining important features and spatial information. The reduced image representations are eventually flattened, where each pixel is then part of a variable vector that might be used to forecast class membership by minimizing a log-loss or 'softmax' function, somewhat analogous to logistic regression for the two-class cases and multinomial regression for the multi-class cases.

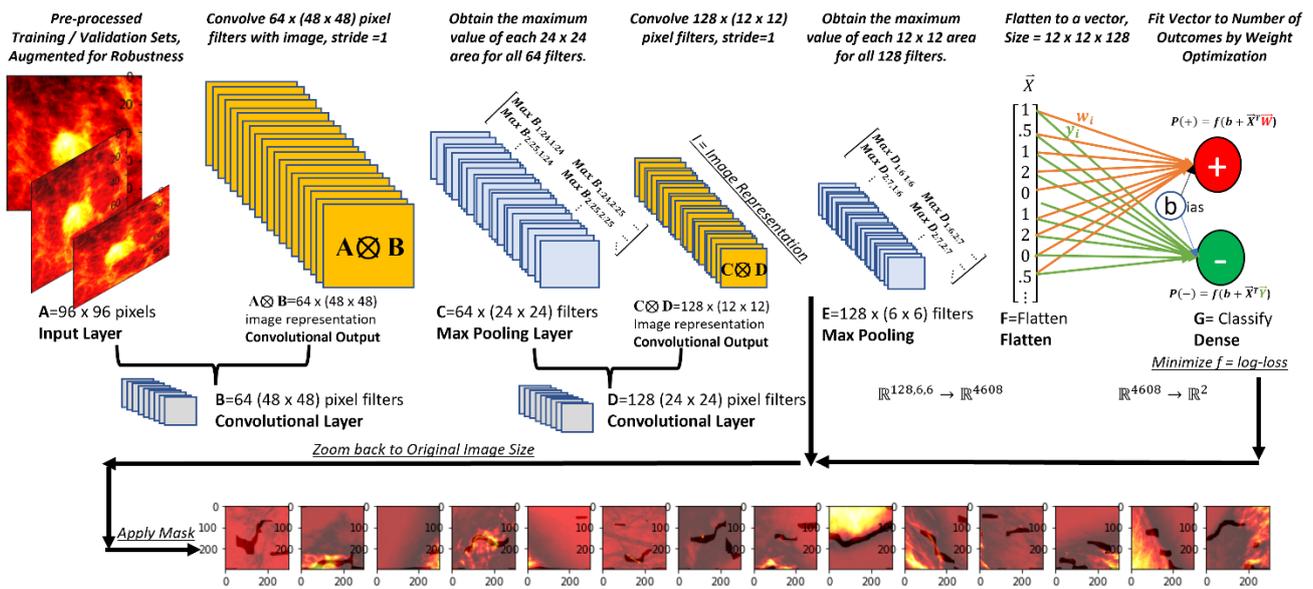


Figure 1. ConvNet architecture and graphical research overview.

2.6. Supervised Classification

Assume that we are evaluating pre-processed images of 96×96 pixels with 1 color channel ('A' in Figure 1). Many image filters of smaller size (e.g., 48×48 pixels, 'B' in Figure 1) are passed over the original image, section by section, one or more pixels at a time. The tensor product, $A \otimes B$, is taken for each of these regions (convolutional output). Moving each of our 48×48 filters over the original 96×96 images 1 pixel at a time generates a new 48×48 matrix. The convolutional layer output, a set of smaller image representations, are additionally processed through a function for classification improvement (a rectifier function that adds additional nonlinearity to images, $f(x) = \max(0, x)$). Smaller filters (24×24 pixels in Figure 1, 'C') are applied to convolutional layers. These filters execute 'maximum pooling' (keeping the maximum value in each region) or 'average pooling' (averaging all values). Multiple layers ('D' and 'E' in Figure 1) are then combined in a network structure. The final image representation is flattened into a single vector (dense layer, 'F' in Figure 1). The inner product of this vector and weights (tuned through nonlinear optimization) are added to a constant (bias) and processed through a 'softmax' function ('G' in Figure 1). The softmax function provides the probability of negative (-) or non-negative (+) imagery Equation (1).

$$P(+)=\frac{e^{-X^TW}}{e^{-X^TW}+e^{-X^TY}}, P(-)=\frac{e^{-X^TY}}{e^{-X^TW}+e^{-X^TY}}. \tag{1}$$

Equation (1) estimates the probability that an image will belong in either the negative (-) or non-negative (+) class. In this equation, X^T is the flattened vector produced by the ConvNet ('F' in Figure 1), and the weights for negative (-) and non-negative (+) conditions are W and Y , respectively. The probability for each class is estimated by dividing the exponent of the weighted vector for each class by the sum of the weighted vectors for both classes.

For two-category classification, only one equation rather than two is needed. A sigmoid or log-loss equation classifies either negative (-) or non-negative (+) studies Equation (2).

$$P(+)=\sigma(x)=\frac{1}{1+e^{-x}} \tag{2}$$

In Equation (2), we define $\sigma(x)$ as the sigmoid function which takes the vector data from the flattened layer and estimates the probability of non-negative (+) results.

ConvNets require tuning of the weights and filter values through nonlinear minimization of a loss function, typically the log-loss (binary cross-entropy, Equation (3)).

$$L = -\frac{1}{N} \sum_{i=1}^N y_i \log(p_i) + (1 - y_i) \log(1 - p_i) \quad (3)$$

In Equation (3), L is the loss function, N is the number of items classified, y is the true class membership (either negative or non-negative), and p is the probability estimate for the group membership. A correctly classified observation (e.g., a true '1' predicted to be such with probability 0.999) results in a value near 0. An incorrectly classified observation (e.g., a true '1' predicted to be '0' with 0.999 probability) results in a value of 4.605. Thus, minimizing binary cross-entropy through nonlinear optimization attempts to improve the classification.

2.7. Specific Architecture, Classification Problem

For this analysis, we evaluated several pre-existing architectures that have shown to perform well on the image classification task. The best-performing architecture of those considered is presented here and is based on the Visual Geometry Group (VGG) of Oxford [35]. The architecture, known as VGG-16, includes 13 convolutional layers and 5 maximum pooling layers. Global average pooling was added to this baseline architecture to further reduce dimensionality, and a single neuron with sigmoid activation function provided the final estimate of the probability of a non-negative result. While we started with pre-trained weights for the VGG-16 architecture, we allowed these to be updated during nonlinear optimization.

Images were batched in groups of fed into the architecture. Validation metrics were estimated, and hyperparameter tuning was performed. The final selected image batch size was 32, and the chosen nonlinear optimizer was an adaptive gradient algorithm known as 'AdaGrad' [36]. A total of 50 epochs were set with early stopping conditions (which activated after epoch 25).

2.8. Unsupervised Region of Interest (ROI) Identification

Traditional radiological imagery is often devoid of ROI segmentation, which depicts the location of anomalies. Without previously identified ROIs (as is the case with the imagery in this study), it is not possible to train supervised models to isolate ROIs. It is, however, possible to return a small image representation from the closest convolutional layer of a model, evaluate the gradient of the loss function with respect to the image at that layer using computational graphs, find the portion of the gradient that maximizes the classification probability, and re-map that gradient location back to the original image. In doing so, the areas of the original image that are most highly influential for the classification model can be found. Further, an image contour overlay can be built to identify the areas most associated with the classification determination based on important values of the gradient of the loss function with respect to the image. In this study, we avoid guided backpropagation as well as other non-gradient based methods based on recent experimental findings [37].

Figure 2 represents the last functions of the ConvNet provided the probability prediction for the image. For each filter F in the last convolutional layer, we seek the gradient of the loss function with respect to the image, $\frac{\partial L}{\partial F}$. The forward pass (green arrows) results in a composite function from the filters through the rectifier, global average pooling, the sigmoid function to the forecast. To calculate $\frac{\partial L}{\partial F}$ requires that we 'backpropagate' the gradients (red arrows) using the chain rule until we are able to estimate our loss gradient with respect to the filter. These calculations are supported by TensorFlow and need not be done manually [38].

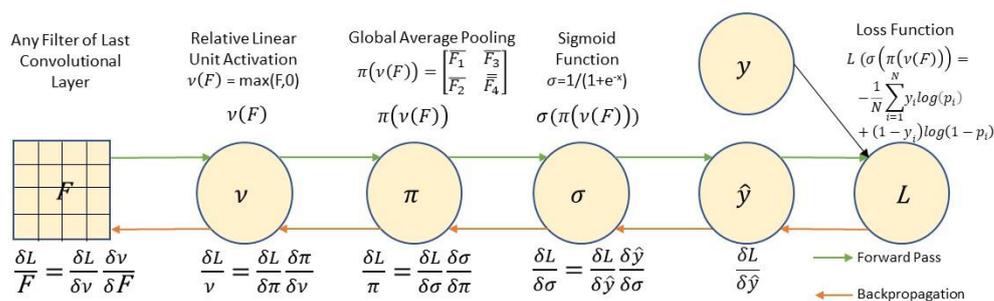


Figure 2. Backpropagation to find $\frac{\delta L}{\delta F}$.

3. Results

3.1. Descriptive Statistics

The data were imbalanced, as 86.96% of the images were negative. ConvNets are robust to imbalanced data, so addressing this imbalance by over or under sampling (for example) was unnecessary. We investigated image augmentation using 10 degrees of random rotation, 20% zoom in and out, 10% height and width shifts, 10% shearing, and horizontal flipping, as augmentation sometimes produces more robust classifiers [39]. We saw no improvement with image augmentation possibly due to issues with local optima and thus proceeded with the original images.

Images were saved as three channel ('RGB') color images to support the baseline VGG-16 architecture. Several of the images were plotted using various color enhancements to verify data quality (see Figure 3 for an example). An autoencoder to reduce image noise was implemented but added no observable benefit to the classification.

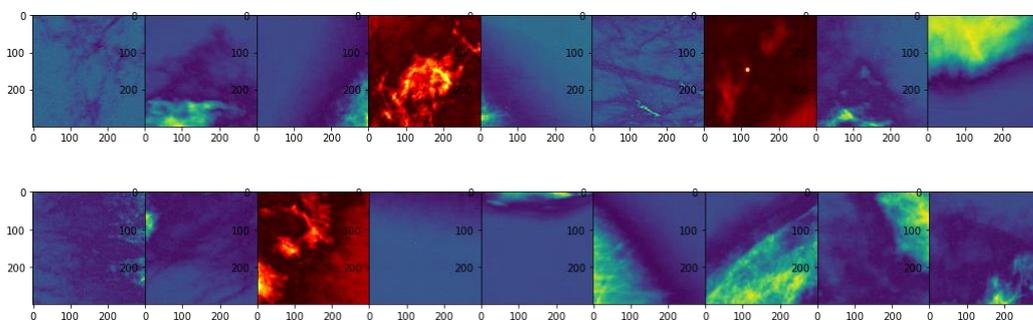


Figure 3. Image plots of slides with non-negative results depicted in red.

3.2. Classification Results

The VGG-16 architecture resulted in 98% classification accuracy, 83% sensitivity (recall of positives), and 99% specificity (recall of negatives) on the pristine test set after 25 epochs. These metrics are all improved over previous studies including [19,20]. The positive predictive value (precision for positive cases) is 0.93, indicating that 93% of the patients informed of positive study results were true positives, much better than [19]. Table 1 provides the complete results of predicting the test result with the trained model, while Table 2 is the confusion matrix (the classification matrix with incorrect and correct predictions labeled).

Table 1 shows that the classification process worked well. The high recall (99%) on the negative patients suggests that overdiagnosis and overtreatment would be reduced. The sensitivity of 83% might even be improved with additional a priori manipulation as well as larger datasets.

In Table 2, the FPR is $(131/1801) = 7.3\%$ while the FNR is $(334/13,563) = 2.5\%$. These metrics are much better than those seen in [18].

Table 1. Metrics for predicted test set data, 97% accuracy.

Metric	Size	Precision	Recall	F1-Score
Negative	13,360	0.98 ⁰	0.99 [*]	0.98 ^{***}
Positive	2004	0.93 ¹	0.83 ^{**}	0.88 ^{***}
Weighted Average	15,364	0.97	0.97	0.97

⁰ Negative Predictive Value, ¹ Positive Predictive Value, ^{*} Sensitivity, ^{**} Specificity, ^{***} Harmonic Mean of Precision & Recall.

Table 2. Confusion matrix for the classification problem.

Actual/Prediction	Negative Prediction	Positive Prediction	Total
Negative	13,229	131	13,360
Positive	334	1670	2004
Total	13,563	1801	15,364

3.3. Unsupervised Gradient Mapping

We used TensorFlow's 'GradientTape' function to calculate the gradient of the loss function at the last convolutional layer's reconstructed image, which includes 512 filters of size 18×18 pixels. After producing the gradient map, we zoomed the image back to 299×299 to produce contour overlays for the images. To further identify areas of concern, we filtered for the top 10% of the image values. In Figure 4, we show the first 10 positive images in the data set in gray scale with the contour mask of the gradient superimposed (showing as pink).

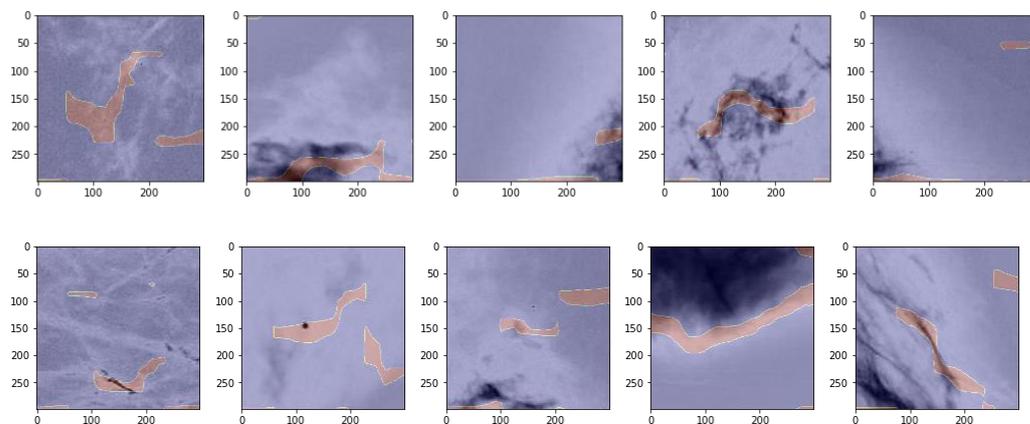
**Figure 4.** The first 10 positive images in the data set with the gradient loss contour mapped.

Figure 4 shows that the unsupervised classification is able to pick up features in the images associated with anomalies. In Row 1, plots 2 through 5 and in Row 2 plots 3 through 5, it is clear that the gradient of the loss with respect to the image considers the darkish areas to be associated with anomalies. In Row 2, plots 1 and 2 appear to identify a high density 'line' and 'point', respectively. The contour plots appear to capture the boundaries of interesting tissue in the plots, which is to be expected, as the maximum values of the gradient of the loss with respect to the image representation are those areas most closely associated with the algorithm's classification decision.

4. Discussion

Supervised classification of mammography detection is effective in achieving high sensitivity, specificity, precision, and overall accuracy. The model presented here was highly specific (99%) with reasonable sensitivity (83%). This particular model would effectively address the problems of overdiagnosis and overtreatment, and slightly address the problems of morbidity and mortality associated with breast cancer. Refinements to the

architecture, weights, and other elements of the model would be expected to improve its performance; however, a relatively common ConvNet architecture was capable of achieving these results on a sufficiently large dataset with reasonably sized images without significant hyperparameter tuning or image manipulation.

Backpropagation techniques worked well to produce contour maps for the loss gradient with respect to the image representation. Zooming these contour maps from the last convolutional layer back to the original image size provided reasonable visual representation of the areas that are (by definition) most closely associated with the classification process. Thus, it is relatively straightforward to determine regions of interest during the ML classification process without running separate models. Providing pre-marked images to physicians, particularly when ML models and physicians have different ‘opinions’ regarding the mammography, is an important consideration for improving FPR and FNR rates.

A limitation of this study is that it uses previously curated images of fixed and relatively small size. Larger files (e.g., Aperio ScanScope.svs image sets) require significantly more preprocessing and computational power. Further, the positive mammography lacks confirmed ROIs, so the contour masks may not be interpreted alone. In addition, this study focused on the classification of non-negative imagery. Classification using the BI-RADS classification scheme is also possible but beyond the scope of this initial study.

Future work involving imagery obtained by the National Cancer Institute under a Data Transfer Agreement (PLCOI 742) is ongoing. This work requires significant processing power from a cloud-based high performance computing cluster supported by Advanced Micro Devices under an in-kind grant.

5. Conclusions

Deep vision holds much promise for breast cancer classification, as it is able to find relationships among spatially related pixels that the human eye cannot detect. Thus, it is possible for deep vision to detect cancers missed by clinical professionals under even optimal conditions assuming that properly classified training data are available from which the algorithms can learn.

ML techniques in general hold great promise for both the supervised classification and unsupervised segmentation problems. These techniques are likely to assist radiologists and clinicians in reducing FNR and FPR rates, addressing the issues of overdiagnosis and overtreatment while simultaneously reducing morbidity and mortality. It is clear that the role of the ML in assisting clinicians with diagnoses and ROI identification will increase given its demonstrated potential.

Author Contributions: Conceptualization, L.F.; methodology, L.F., N.B., C.P.F., A.M. and D.D.; validation, N.B. and C.P.F.; writing, L.F., N.B., C.P.F., A.M. and D.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research is supported by a funded research leave grant awarded by Texas State University.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are available online with a free account. The address for access follows: <https://www.kaggle.com/skooch/ddsm-mammography>, accessed on 5 January 2021.

Acknowledgments: Special acknowledgement is made to Advanced Micro Devices (AMD) who provided an in-kind grant for computational support.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA A Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]

2. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2020. *CA A Cancer J. Clin.* **2020**, *70*, 7–30. [[CrossRef](#)]
3. WHO. Breast Cancer. Available online: <https://www.who.int/news-room/fact-sheets/detail/breast-cancer> (accessed on 18 October 2021).
4. CDMRP. Breast Cancer Landscape. Available online: <https://cdmrp.army.mil/bcrp/pdfs/Breast%20Cancer%20Landscape2020.pdf> (accessed on 1 March 2021).
5. Zhu, K.; Devesa, S.S.; Wu, H.; Zahm, S.H.; Jatoi, I.; Anderson, W.F.; Peoples, G.E.; Maxwell, L.G.; Granger, E.; Potter, J.F. Cancer incidence in the US military population: Comparison with rates from the SEER program. *Cancer Epidemiol. Prev. Biomark.* **2009**, *18*, 1740–1745. [[CrossRef](#)] [[PubMed](#)]
6. Cancer.org. How Common is Breast Cancer? Available online: <https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html> (accessed on 18 October 2021).
7. Ross, R.K.; Paganini-Hill, A.; Wan, P.C.; Pike, M.C. Effect of hormone replacement therapy on breast cancer risk: Estrogen versus estrogen plus progestin. *J. Natl. Cancer Inst.* **2000**, *92*, 328–332. [[CrossRef](#)]
8. Plasticsurgery.org. National Plastic Surgery Statistics. 2018. Available online: <https://www.plasticsurgery.org/documents/News/Statistics/2018/plastic-surgery-statistics-report-2018.pdf> (accessed on 18 October 2021).
9. Tripodi, D.; Amabile, M.I.; Varanese, M.; D’Andrea, V.; Sorrenti, S.; Cannistrà, C. Large cell anaplastic lymphoma associated with breast implant: A rare case report presentation and discussion of possible management. *Gland Surg.* **2021**, *10*, 2076–2080. [[CrossRef](#)] [[PubMed](#)]
10. De Jong, W.H.; Panagiotakos, D.; Proykova, A.; Samaras, T.; Clemens, M.W.; De Jong, D.; Hopper, I.; Rakhorst, H.A.; Santanelli di Pompeo, F.; Turner, S.D.; et al. Final opinion on the safety of breast implants in relation to anaplastic large cell lymphoma: Report of the scientific committee on health, emerging and environmental risks (SCHEER). *Regul. Toxicol. Pharmacol.* **2021**, *125*, 104982. [[CrossRef](#)] [[PubMed](#)]
11. McCarthy, B.D.; Yood, M.U.; MacWilliam, C.H.; Lee, M.J. Screening Mammography Use: The Importance of a Population Perspective. *Am. J. Prev. Med.* **1996**, *12*, 91–95. [[CrossRef](#)]
12. Witten, M.; Parker, C.C. Screening mammography: Recommendations and controversies. *Surg. Clin.* **2018**, *98*, 667–675.
13. Duffy, S.W.; Tabár, L.; Yen, A.M.F.; Dean, P.B.; Smith, R.A.; Jonsson, H.; Törnberg, S.; Chen, S.L.S.; Chiu, S.Y.H.; Fann, J.C.Y. Mammography screening reduces rates of advanced and fatal breast cancers: Results in 549,091 women. *Cancer* **2020**, *126*, 2971–2979. [[CrossRef](#)]
14. Levinsohn, E.; Altman, M.; Chagpar, A.B. Article Commentary: Controversies Regarding the Diagnosis and Management of Ductal Carcinoma in Situ. *Am. Surg.* **2018**, *84*, 1–6. [[CrossRef](#)]
15. Nelson, H.D.; O’Meara, E.S.; Kerlikowske, K.; Balch, S.; Miglioretti, D. Factors associated with rates of false-positive and false-negative results from digital mammography screening: An analysis of registry data. *Ann. Intern. Med.* **2016**, *164*, 226–235. [[CrossRef](#)]
16. Liu, X.; Tang, J. Mass Classification in Mammograms Using Selected Geometry and Texture Features, and a New SVM-Based Feature Selection Method. *IEEE Syst. J.* **2014**, *8*, 910–920. [[CrossRef](#)]
17. Nishikawa, R.M.; Bae, K.T. Importance of better human-computer interaction in the era of deep learning: Mammography computer-aided diagnosis as a use case. *J. Am. Coll. Radiol.* **2018**, *15*, 49–52. [[CrossRef](#)]
18. Saki, F.; Tahmasbi, A.; Soltanian-Zadeh, H.; Shokouhi, S.B. Fast opposite weight learning rules with application in breast cancer diagnosis. *Comput. Biol. Med.* **2013**, *43*, 32–41. [[CrossRef](#)] [[PubMed](#)]
19. Ertosun, M.G.; Rubin, D.L. Probabilistic visual search for masses within mammography images using deep learning. In Proceedings of the 2015 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Washington, DC, USA, 9–12 November 2015; pp. 1310–1315.
20. Muramatsu, C.; Higuchi, S.; Morita, T.; Oiwa, M.; Kawasaki, T.; Fujita, H. Retrieval of reference images of breast masses on mammograms by similarity space modeling. In Proceedings of the 14th International Workshop on Breast Imaging (IWBI 2018), Atlanta, GA, USA, 8–11 July 2018; p. 1071809.
21. Salama, W.M.; Aly, M.H. Deep learning in mammography images segmentation and classification: Automated CNN approach. *Alex. Eng. J.* **2021**, *60*, 4701–4709. [[CrossRef](#)]
22. Muhammad, M.; Zeebaree, D.; Brifcani, A.M.A.; Saeed, J.; Zebari, D.A. Region of Interest Segmentation Based on Clustering Techniques for Breast Cancer Ultrasound Images: A Review. *J. Appl. Sci. Technol. Trends* **2020**, *1*, 78–91. [[CrossRef](#)]
23. Drukker, K.; Giger, M.L.; Horsch, K.; Kupinski, M.A.; Vyborny, C.J.; Mendelson, E.B. Computerized lesion detection on breast ultrasound. *Med. Phys.* **2002**, *29*, 1438–1446. [[CrossRef](#)]
24. Preim, B.; Botha, C.P. *Visual Computing for Medicine: Theory, Algorithms, and Applications*; Morgan Kaufman (Elsevier): Waltham, MA, USA, 2013.
25. Huang, Q.-H.; Lee, S.-Y.; Liu, L.-Z.; Lu, M.-H.; Jin, L.-W.; Li, A.-H. A robust graph-based segmentation method for breast tumors in ultrasound images. *Ultrasonics* **2012**, *52*, 266–275. [[CrossRef](#)]
26. Çiğla, C.; Alatan, A.A. Efficient graph-based image segmentation via speeded-up turbo pixels. In Proceedings of the 2010 IEEE International Conference on Image Processing, Piscataway, NJ, USA, 14–19 March 2010; pp. 3013–3016.
27. Zeiler, M.D.; Fergus, R. Visualizing and understanding convolutional networks. In *European Conference on Computer Vision*; Springer: Cham, Switzerland, 2014; pp. 818–833.

28. Springenberg, J.T.; Dosovitskiy, A.; Brox, T.; Riedmiller, M. Striving for simplicity: The all convolutional net. *arXiv* **2014**, arXiv:1412.6806.
29. Dwaraknath, A.; Menghani, D.; Mongia, M. Fast Unsupervised Object Localization. 2016. Available online: http://vision.stanford.edu/teaching/cs231n/reports/2016/pdfs/285_Report.pdf (accessed on 1 October 2021).
30. Heath, M.; Bowyer, K.; Kopans, D.; Moore, R.; Kegelmeyer, W. The digital database for screening mammography. In *Proceedings of the Fifth International Workshop on Digital Mammography*; Yaffe, M.J., Ed.; Medical Physics Publishing: Madison, WI, USA, 2001; pp. 212–218.
31. Lee, R.S.; Gimenez, F.; Hoogi, A.; Miyake, K.K.; Gorovoy, M.; Rubin, D.L. A curated mammography data set for use in computer-aided detection and diagnosis research. *Sci. Data* **2017**, *4*, 1–9. [[CrossRef](#)]
32. Scuccimara, E. DDSM Mammography. Available online: <https://www.kaggle.com/skooch/ddsm-mammography> (accessed on 5 October 2021).
33. Fulton, L.V. Breast Cancer. Available online: <https://github.com/dustoff06/BreastCancers> (accessed on 5 October 2021).
34. Yue, H.; Liu, J.; Yang, J.; Sun, X.; Nguyen, T.Q.; Wu, F. Ienet: Internal and external patch matching convnet for web image guided denoising. *IEEE Trans. Circuits Syst. Video Technol.* **2019**, *30*, 3928–3942. [[CrossRef](#)]
35. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
36. Lydia, A.; Francis, S. Adagrad—An optimizer for stochastic gradient descent. *Int. J. Inf. Comput. Sci.* **2019**, *6*, 566–568.
37. Adebayo, J.; Gilmer, J.; Muelly, M.; Goodfellow, I.; Hardt, M.; Kim, B. Sanity checks for saliency maps. *arXiv* **2018**, arXiv:1810.03292.
38. TensorFlow. tf.GradientTape. Available online: https://www.tensorflow.org/api_docs/python/tf/GradientTape (accessed on 5 October 2021).
39. Hernández-García, A.; König, P. Further advantages of data augmentation on convolutional neural networks. In *Proceedings of the International Conference on Artificial Neural Networks*; Springer: Cham, Switzerland, 2018; pp. 95–103.