

LITHOGRAPHY HOTSPOT DETECTION USING VISION TRANSFORMER

Latha PH¹, Jayanth P², Ali Akbar P³

¹Head of Dept of ISE, Sambhram Institute of Technology, Bengaluru Urban, Karnataka, India.

^{2,3}Final Year Students, Dept. Of ISE, Sambhram Institute of Technology, Bengaluru Urban, Karnataka, India.

DOI: <https://www.doi.org/10.58257/IJPREMS33068>

ABSTRACT

In the process of IC design, lithography can be defined as the process of reissuing the pattern of the mask on a Silicon wafer. Lithography is an essential step in this process as it enables point size to drop which further helps in dwindling device size. This nonstop drop in point size may lead to printability issues and hotspots. Presence of hotspots can beget the circuit to fail, so it's veritably important to descry these hotspots with high delicacy. preliminarily colorful simulation, machine leaning and deep literacy grounded ways have been enforced to break this issue. In this paper, a system to identify hotspots using Vision Mills is proposed. Other deep literacy ways, similar as CNNs and ANNs have also been used for comparison purposes. All three ways are enforced on five datasets. ViT gives an overall average delicacy of 98.05 which is 1.39 advanced than delicacy of CNNs and 2.04 advanced than delicacy given by ANNs. Although the ViTs prove the stylish in terms of overall delicacy, but at dataset position its performance can be bettered. Three out of five datasets have delicacy advanced than 99 and for rest two it's slightly above 95. In future, we wish to ameliorate delicacy for these two datasets by perfecting the model and reducing imbalance in the datasets.

Keywords: Deep literacy, Lithography, Hotspot Detection, Vision Transformer, complication Neural Network, Artificial Neural Network.

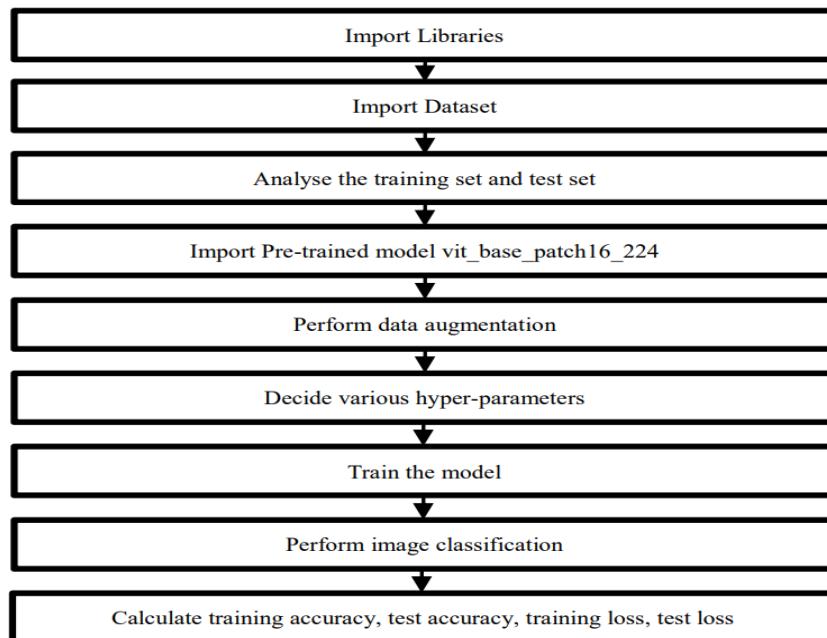
1. INTRODUCTION

In the process of IC fabrication, patterns are generated on a silicon wafer. These patterns are first obtained on a mask and then transferred on silicon wafer through the process known as lithography. In order to fit more and more transistors in the same area, feature size needs to get smaller. In optical lithography, feature size is directly proportional to wavelength. Mathematically,

$$F = C \cdot \lambda / n$$

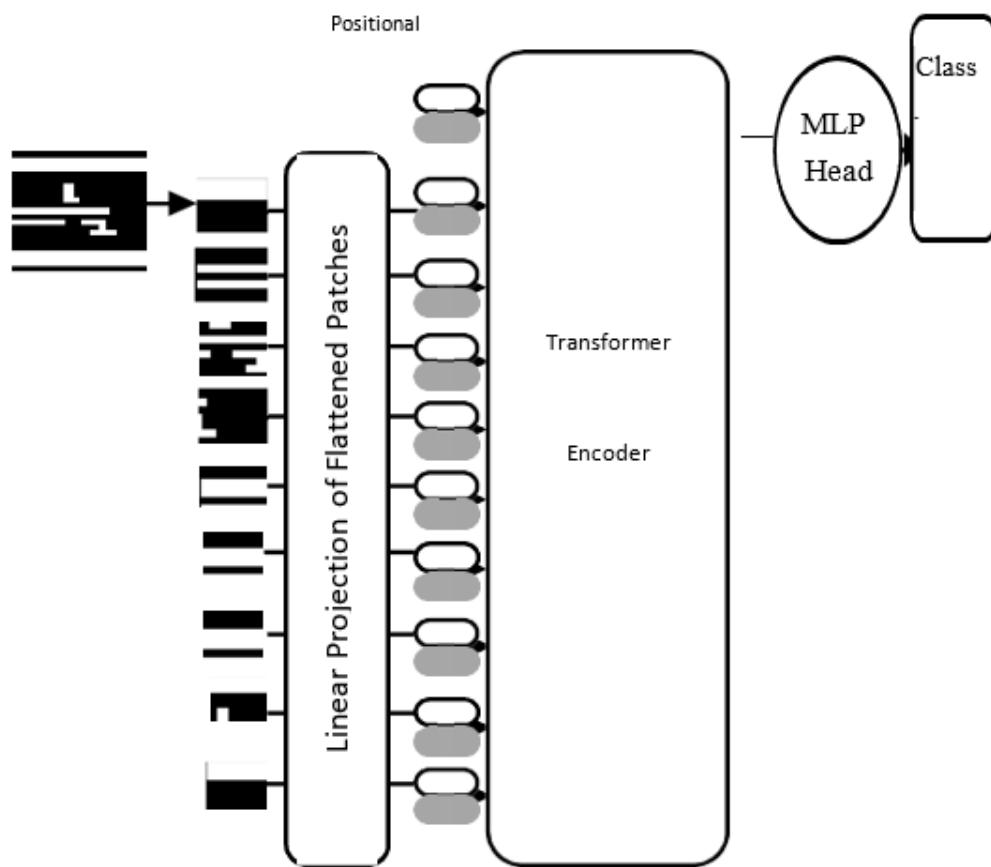
Here, F is the feature size, C is the Rayleigh constant which measures how difficult lithography is, λ is the wavelength, and n is the numerical aperture. The most effective way to reduce feature size is to reduce the wavelength of light. This wavelength reduction leads to printability problems and degradation in resolution. Although Resolution Enhancement Techniques such as Optical Proximity Correction, Sub Resolution Assist Feature (SRAF) are employed to improve the process, but at some locations, differences exist between patterns on mask and wafer. Positions where patterns have dimensions more or less than the defined threshold are known as hotspots. In electron beam lithography, electrons scatter, and these scattered electrons may cover a different path than the one drawn in mask, which may cause hotspots. The hotspots may lead to an open circuit or short circuit; hence detecting them is very important. Various Simulations, Pattern Matching, Machine Learning, and Deep Learning based techniques have been implemented to eliminate of hotspots. Simulation and Deep Learning based techniques are very expensive in terms of computational time and complexity. Although Pattern Matching techniques are faster, they fail to detect previously unseen hotspots. Machine Learning based techniques are faster and more efficient in detecting previously unseen hotspots, but due to an imbalance in dataset, false positives remain a big problem in this method, which leads to more time for preprocessing the data. In this work, we propose using Vision Transformer (ViT) technique to identify lithography hotspots. Transformers have a wide range of applications in the field of Natural Language Processing. However, these are only used to a small extent in Computer Vision based applications due to high number of computations involved, which are not possible to achieve with hardware. ViT aims to overcome these limitations by converting image into patches, passing them through the transformer encoder structure and finally classifying them. This technique, introduced by Alexey Dosovitskiy et al. in June 2021 has not been previously utilized for detecting hotspots. Section 4 discusses details of the datasets and all experiments performed for implementation and comparison of above mentioned techniques. From these trials, it can be observed that ViT gives 1.39% higher overall average accuracy than the accuracy of CNNs and 2.04% higher than the accuracy given by ANNs. While comparing to already existing works, ViT performs the best or comparable to best for three out of five datasets, but it is not able to supplant all the existing methods for all the datasets. In section 5, results are shown, followed by conclusions and future scope.

2. METHODOLOGY



1. Importing Libraries: Libraries such as PyTorch and torchvision are essential for building and training neural network models for computer vision tasks. PyTorch provides a flexible and efficient framework for tensor computations and automatic differentiation, while torchvision offers useful utilities and pre-trained models for working with image data.
2. Import Dataset: In machine learning and deep learning projects, datasets are essential for training and evaluating models. Datasets typically consist of input samples (images, text, etc.) and corresponding labels (classifications, categories, etc.). In this context, we import the dataset using PyTorch's ImageFolder class, which organizes images into folders based on their class labels.
3. Analysing the Training Set and Test Set: Understanding the characteristics of the training and test sets is crucial for model development and evaluation. Analysis may include examining the distribution of classes, the number of samples per class, and visualizing sample images. This analysis helps identify potential biases and informs decisions during model training and evaluation.
4. Importing Pre-trained Model (ViT): Vision Transformer (ViT) is a deep learning architecture that applies the transformer architecture, originally developed for natural language processing tasks, to computer vision tasks. Pre-trained ViT models, such as vit_base_patch16_224, have been trained on large-scale image datasets (e.g., ImageNet) and can be fine-tuned for specific vision tasks, such as image classification or object detection.
5. Performing Data Augmentation: Data augmentation is a technique used to increase the diversity of training data by applying random transformations such as cropping, rotation, flipping, and scaling. Augmentation helps the model generalize better to unseen data and reduces overfitting by exposing the model to variations in the input data distribution.
6. Deciding Various Hyperparameters: Hyperparameters are parameters that govern the training process and model architecture but are not learned from the data. Examples include learning rate, batch size, number of epochs, and optimizer settings. Deciding on appropriate hyperparameters requires experimentation and domain knowledge to balance model performance and computational resources.
7. Training the Model: Training a neural network involves optimizing its parameters (weights and biases) to minimize a predefined loss function. This is typically done using an optimization algorithm such as stochastic gradient descent (SGD) or its variants (e.g., Adam). During training, the model iteratively updates its parameters based on gradients computed from the training data until convergence or a predefined stopping criterion.
8. Performing Image Classification: Image classification is a computer vision task where the goal is to assign a label or category to an input image based on its content. In this context, the trained ViT model is used to predict the class labels for images in the test set.
9. Calculating Training and Test Metrics: Training accuracy, test accuracy, training loss, and test loss are common metrics used to evaluate the performance of machine learning models. Training accuracy and loss measure the model's performance on the training data, while test accuracy and loss assess its generalization ability on unseen test data.

3. MODELING AND ANALYSIS



In transformers, one can pay attention to the things which are far away; this is not possible in CNNs. One feature that has been observed with ViTs is that they perform better than ResNETs only when training data is sufficiently large; otherwise they are equally or less effective. In CNNs, integration is done over a pixel, which connects to its neighborhood. Then that neighborhood connects to its neighborhood and so on. This is known as local attention. ViTs work on the principle of global attention i.e., all the points are connected at once

4. RESULTS AND DISCUSSION

In this Section effects and deliberation of the study is penned. They may also be broken up into subsets with short, discovering captions. This section should be compartmented in character size 10 pt Times New Roman.

Dataset	Accuracy					Average Accuracy
	Dataset 1	Dataset 2	Dataset 3	Dataset 4	Dataset 5	
ViT	95.48	99.37	95.77	99.83	99.8	98.05
CNNs	94.37	98.81	90.91	99.45	99.79	95. 96.66
ANNs	89.58	97.73	94.58	96.68	99.48	96.01

Fig. 4 shows that for all the datasets ViT gives the best results

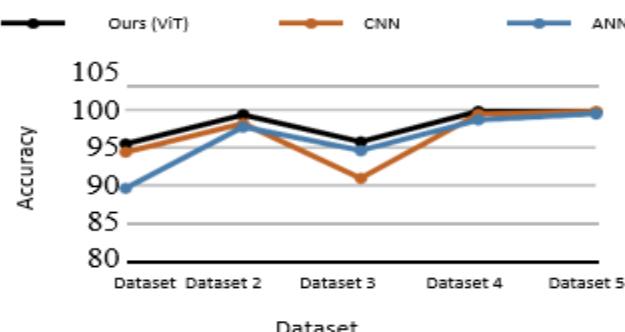


Fig. 4. Comparing accuracies for ViT, CNNs and ANNs

CNNs perform moderately well for datasets 1, 2, 4 and 5 and worst for dataset 3. ANNs perform poorest for sub-datasets 1, 2, 4, 5 and moderately well for sub-dataset 3. In terms of overall accuracy ViTs give 1.39% better accuracy than CNNs and 2.04% better accuracy than ANNs.

6. CONCLUSION

In this paper, lithography hotspots have been detected using Vision Transformers. To see if this proposed technique gives better results than the already existing deep learning techniques, we applied CNNs and ANNs to solve this problem along with ViT. Table 2, shows that in terms of overall accuracy ViT gives 1.39% better accuracy than CNNs and 2.04% better accuracy than ANNs. Considering individual dataset wise accuracies, ViT performs better or as well as CNNs for each dataset. For three out of five datasets, accuracy on the test set is more than 99%, and for the other two, it is more than 95%. Table 3 shows comparison of ViT model with existing research works, and it can be seen that in terms of overall accuracy, ViT gives the best results. At the individual dataset level for three out of five datasets, it provides the best or comparable results but lags for two datasets. From the results, it can be concluded that although the proposed technique performs better than many already existing state of the art techniques, it can only supplant some.

7. REFERENCES

- [1] Sivakumar, S. EUV lithography: Prospects and challenges In 16th Asia and South Pacific Design Automation Conference, 2011, pp. 402-402 doi: 10.1109/ASPDAC.2011.5722221
- [2] 2. Robles, J. A.T.; Mostafa, S.; Madkour, K. & Wuu, J.Y. Hotspot Detection Based on Machine Learning. In United States Patent Application Publication, 2013, , [Online], Available: <https://patents.google.com/patent/US20130031522>. [Accessed: June 2021]
- [3] Gkelios, S.; Boutalis, Y. & Chatzichristofis, S.A. Investigating the Vision Transformer Model for Image Retrieval Tasks. 11 Jan 2021, [Online], Available: arXiv:2101.03771v1. [Accessed: June 2021]
- [4] Ye, W.; Lin , Y.; Li, M.; Liu, Q.; & Pan, D.Z. LithoROC: lithography hotspot detection with explicit ROC optimization. In ASPDAC '19: Proceedings of the 24th Asia and South Pacific Design Automation Conference, January 2019, pp. 292–298
- [5] Tomioka, Y.; Matsunawa, T.; Kodama, C. & Nojima, S. Lithography hotspot detection by two-stage cascade classifier using histogram of oriented light propagation. In 22nd Asia and South Pacific Design Automation Conference (ASP-DAC), 2017, pp. 81-86 doi: 10.1109/ASPDAC.2017.7858300.
- [6] u, Y.; Chan, Y.; Sinha, S.; Jiang, I. H. & Chiang, C. Accurate processhotspot detection using critical design rule extraction In DAC Design Automation Conference 2012, 2012, pp. 1163-1168.1
- [7] Yu, Y.; Chan, Y.; Sinha, S.; Jiang, I. H. & Chiang, C. Accurate processhotspot detection using critical design rule extraction In DAC Design Automation Conference 2012, 2012, pp. 1163-1168
- [8] Reddy, G. R.; Xanthopoulos, C. & Makris, Y. Enhanced hotspot detection through synthetic pattern generation and design of experiments. In IEEE 36th VLSI Test Symposium (VTS), 2018, pp. 1- 6 doi: 10.1109/VTS.2018.8368646.
- [9] Gao, J.; Yu, B.; Ding, D. & Pan, D. Z. Lithography hotspot detection and mitigation in nanometer VLSI. In IEEE 10th International Conference on ASIC, 2013, pp. 1-4 doi:10.1109/ASICON.2013.6811917.
- [10] Kim, S.K.; Lee, J.E.; Park, S.W. & Oh, H.K. Optical lithography simulation for the whole resist process. In Current Applied Physics, Volume 6, Issue 1, 2006, pp. 48-53 doi.: 10.1016/j.cap.2004.12.003.
- [11] Pan, D. Z. Lithography-aware physical design. In 6th International Conference on ASIC, 2005, pp. 1172-1173 doi:10.1109/ICASIC.2005.1611242.
- [12] Alexey Dosovitskiy, Lucas Beyer Et al.: An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale, June 2021 [Online], Available : <https://doi.org/10.48550/arXiv.2010.11929> [Accessed: May 2022]
- [13] Zhang, H.; Zhu, F.; Li, H.; Young, E.F.Y.; Yu, B. Bilinear Lithography Hotspot Detection Share. In ISPD '17: Proceedings of the 2017 ACM on International Symposium on Physical Design, March 2017, Pp. 7– 14 doi.: 10.1145/3036669.3036673.