LITHOGRAPHY HOTSPOT DETECTION USING VISION TRANSFORMER

Project Report
Submitted in Partial Fulfilment of the Requirements
For the Award of the Degree
Of

Master of Technology
In
Signal Processing and Digital Design

by

Sumedha
2K19/SPD/17

Under the supervision of

Prof. Rajesh Rohilla

Department of Electronics and Communication Engineering

Delhi Technological University

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

July 2021

Department of Electronics and Communication Engineering

Delhi Technological University

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Sumedha, Roll No. 2K19/SPD/17 student of M.Tech (Signal Processing and Digital Design), hereby declare that the project titled "**Lithography Hotspot Detection using Vision Transformer**" which is submitted by me to the Department of Electronics and Communication Engineering, Delhi Technological University, Delhi in partial fulfilment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any Degree, Diploma Associate ship, Fellowship or other similar title or recognition.

Sumedha

Place: Delhi — **Sumedha**

Date: 31/07/2021 — 2K19/SPD/17

Department of Electronics and Communication Engineering

Delhi Technological University

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Project Dissertation titled "**Lithography Hotspot Detection using Vision Transformer**" which is submitted by Sumedha, Roll No. 2K19/SPD/17, Electronics and Communication Engineering Department, Delhi Technological University, Delhi, in partial fulfilment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the students under my supervision. To the best of my knowledge this work has not been submitted in part or full for any Degree to this University or elsewhere.

Place: Delhi **PROF. RAJESH ROHILLA**

Date: 31/07/2021 SUPERVISOR

ABSTRACT

In the process of IC design, lithography can be defined as the process of reprinting the pattern of mask on Silicon wafer. Lithography is one of the most important steps in the process as it enables Moore's law to be satisfied, for this feature size needs to be decreased every couple of years. This continuous decrease in feature size may lead to printability issues and hence hotspots. Presence of hotspots can lead to complete failure of the circuit, so it is very important to detect these hotspots with high accuracy. Previously various simulation, machine leaning and deep learning based techniques have been implemented to solve this problem. In this work, we propose a method to identify hotspots using Vision Transformers. Along with this, we also use other deep learning techniques such as CNNs and ANNs for comparison purposes. ViTs give an overall accuracy of 98.05% which is 1.39% higher than accuracy of CNNs and 2.04% better accuracy of ANNs. Although the ViTs prove the best in terms of overall accuracy, but at sub-dataset level its performance can be improved. Two out of five sub-datasets have accuracy slightly above 95% and for rest three it is above 99%. In future, we wish to improve accuracy for these two sub-datasets by improving our model and reducing imbalance in the sub-datasets.

ACKNOWLEDGEMENT

I would like to express my gratitude towards all the people who have contributed their precious time and effort to help me without whom it would not have been possible for me to understand and complete the project.

With profound sense of gratitude I would like to thank Prof. Rajesh Rohilla, my Project guide, for his guidance, patience, support and encouragement throughout the period this work was carried out. His readiness for consultation at all times, his educative comments, his concern and assistance even with practical things have been invaluable.

Sumedha
2K19/SPD/17

CONTENTS

List of Tables

List of Figures

CHAPTER 1

INTRODUCTION

In the process of IC fabrication we wish to generate patterns on silicon wafer. These patterns are first obtained on a mask and then transferred on silicon wafer, through the process known as lithography [43]. Lithography is the engine that derives Moore's law. From lithography point of view, in order to make fit more and more transistors in the same area, size of transistor needs to get smaller. Hence, in order to continuously follow the trends of Moore's law, patterning solutions need to be developed in cost effective way. In optical lithography feature size is directly proportional to wavelength. Mathematically,

$$f = \frac{C \cdot \lambda}{n} \tag{1.1}$$

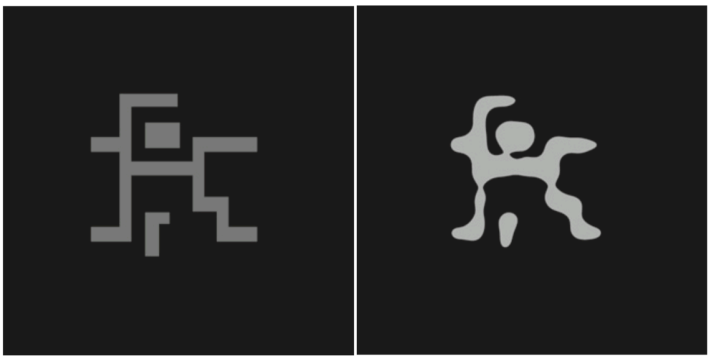
Here, f is the feature size,

C is Rayleigh constant which measures how difficult lithography is,

λ is the wave length,

and n is the numerical aperture.

The most effective way to reduce feature size is to reduce the wavelength of light. Some design rules must be followed by the layout being transferred on the substrate such as threshold value of edge widths of a pattern, minimum spacing between two patterns etc. should remain same as mentioned. This reduction of wavelength leads to printability problems and degradation in resolution [46]. Fig. 1.1 shows this phenomenon.



a). b).

Fig. 1.1 a). Pattern on the mask b). Pattern on the wafer [36]

Many Resolution Enhancement Techniques such as Optical Proximity Correction etc. are employed in order to improve the process, but still at some locations differences exist between pattern on mask and wafer [43]. Positions where patterns have dimensions more or less than the defined threshold are known as hotspots [16]. In electron beam lithography, electrons scatter and these scattered electrons may cover a different path than the one drawn in mask. This may be a cause of hotspots in electron beam lithography [43]. These hotspots may lead to circuit failure, hence detecting them is very important. Fig. 1.2 shows the hotpot pattern.

Fig. 1.2 Hotspot Pattern

Various Simulations, Pattern Matching, Machine Learning and Deep Learning based techniques have been implemented to get rid of hotspots. These techniques have been discussed in Chapter 2. In this work, we propose a new technique called Vision Transformer (ViT) for detecting lithography hotspots. ViT converts images into patches and then passes them through transformer in order to classify them [2]. ViT and lithography are explained in detail in Chapter 3. Chapter 4 discusses about datasets and experiments performed. Chapter 5 results are shown followed by conclusions and future scope in chapter 6.

CHAPTER 2

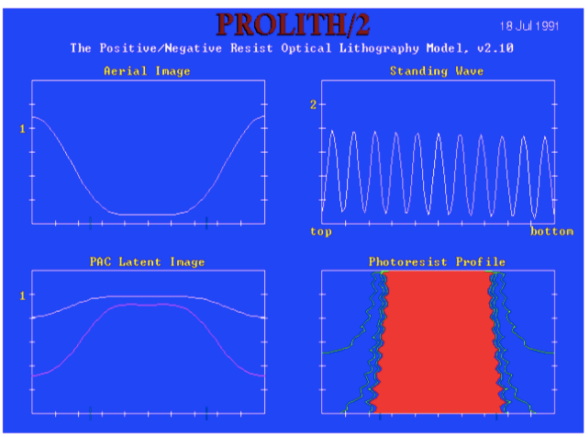
LITERATURE REVIEW

For identification of lithography hotspots first of all various Resolution Enhancement Techniques (RETs) such as Optical Proximity Correction (OPC) and Sub Resolution Assist Feature (SRAF) are applied [16]. After which different models can be applied, which can be divided on two basis: based on methods used for detecting lithography hotspots and based on evaluation parameters.

2.1 Based on methods used for detecting lithography hotspots:

Based on methods used for detecting lithography hotspots, all the models can be divided in five categories, these being: Simulation based methods, Pattern Matching, Machine Learning based Methods, Deep Learning based methods and Lithography Hotspot Mitigation.

In 1979, first optical lithography simulation method called SAMPLE was introduced. This technique provided better results for grids with greater size. In 1985 another method for simulation called The Positive Resist Optical Lithography (PROLITH) was introduced, which made the process of lithography highly accessible as it was the first time when a model could run on a Personal Computer. Various improved versions of PROLITH are still used as simulator for optical lithography process. Simulators used for electron beam are electromagnetic field simulator ProMAX, Monte Carlo, ProBEAM [29]. Full layout simulation is highly accurate way for recognizing hotspots, but it is very expensive in terms of computation time and complexity [7].



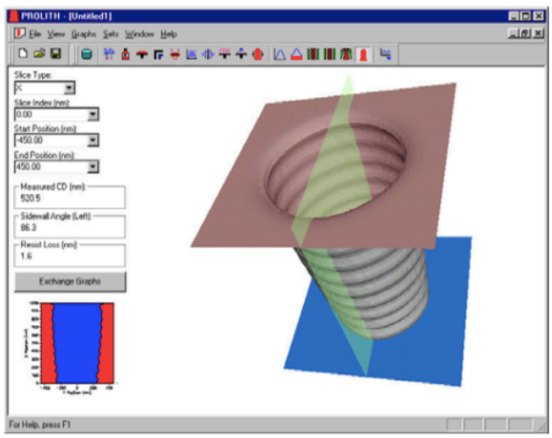


Fig. 2.1 PROLITH: (a) the DOS version 2.1, circa 1991, and (b) the Windows version 6.0, circa 1999 [29]

To know if some place is a hotspot or not, some design rules must be followed. These designs define the minimum distance between two pattens on the same mask, so that they don't overlap, in case of same pattern they define gap between edges etc. Fig represents some basic design rules [25].

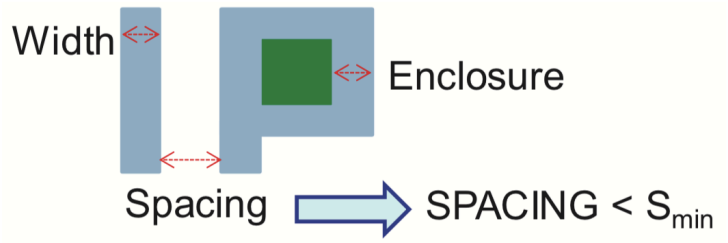


Fig. 2.2 Basic Design Rules for Lithography [25]

Pattern Matching algorithms like string search, tangent space, transitive closure graph, template matching and dual graph first check if each image of the dataset provided follows these rules or not, then a search algorithm is applied to look for hotspots [9]. Graph based techniques create dual or transitive graph for the layouts provided, layouts with edges or adjacent patterns having lesser spacing or greater width etc. than mentioned in rules are considered as hotspots. In this process the chance of non-hotspots being detected as hotspots is large [25]. Template matching methods move the pattern to be detected over entire mask pixel wise and pattern where some disruptions from mask are seen are termed as faulty or with hotspot [21]. The string based methods covert layouts that are two dimensional into a single dimension, these single dimensional structures are named strings. Then, search operations using distance arrays are done to find strings with hotspots [21, 28, 36]. Although the pattern matching based methods are faster than simulations but they fail to detect previously unseen hotspots.

In some machine and deep learning models features need to be extracted before performing classification in order to reduce the size of training data, hence leading to increase in speed. Many algorithms like Topological Classification combined with Critical feature extraction, Concentric Circle Area Sampling (CCAS), Density Based Layout Feature (DBLF),

Histogram Oriented Light Propagation (HOLP), Maximal Circle Mutual Information (MCMI), Encoder-Decoder based, Matrix based Concentric Circle Sampling (MCCS) etc. extract feature from the dataset and check them if they follow design rules or not, and then classification process takes place. DBLF technique divides the layout into sub-regions and calculates their densities as shown in fig. This information is represented in form of vectors and then can be classified using machine learning [14]. DBLF based models classify layouts with same density as shown in fig. in same class, which may cause errors [7].

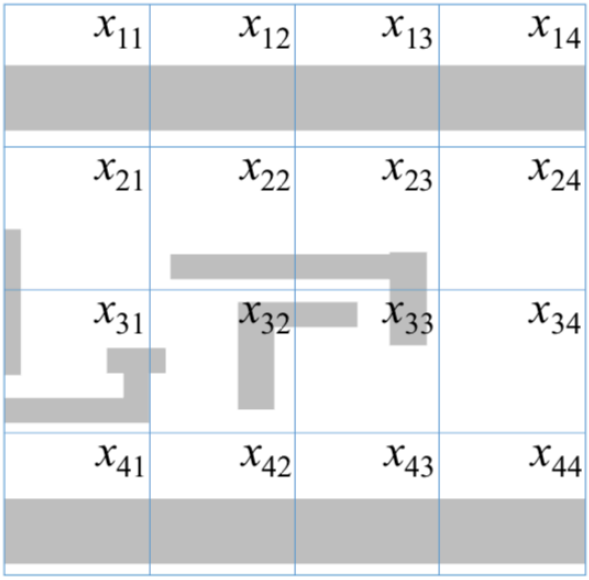


Fig. 2.3 DBLF [14]

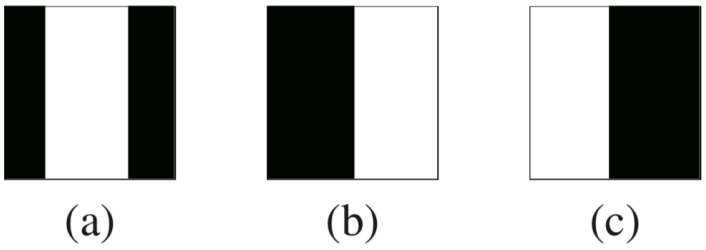


Fig. 2.4 DBLF same density [7]

To overcome weakness of DBLF, HLOP is used, which along with density can also capture direction in which light transmits by calculating Histogram Oriented Gradient (HOG) of regions obtained after performing Gaussian Blurring on image. HLOP process can be understood from following fig. [7].

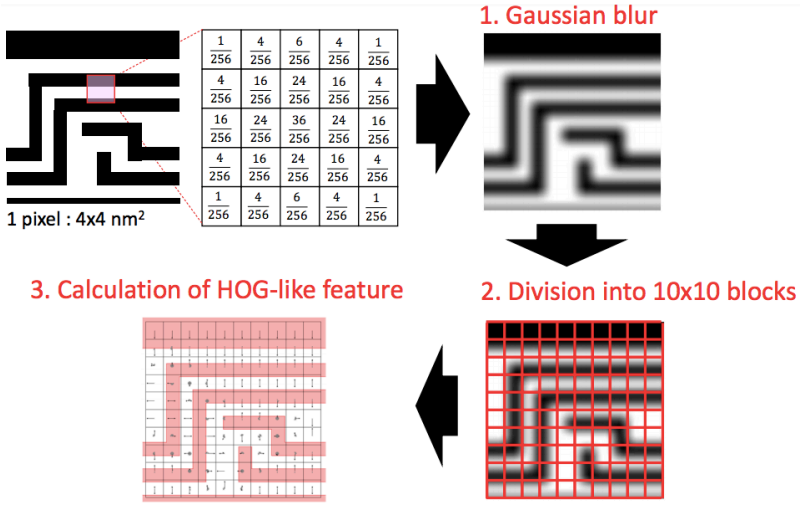


Fig. 2.5 HOLP [7]

Topological classification combined with Critical feature extraction is based on deriving geometry based i.e topological (preserved) and process based i.e. non-topological feature. According to different topologies like distance between edges of a pattern, distance between two patterns etc. and non-topologies like number of corners, number of points touching etc. clusters are made and an SVM kernel related to each critical feature is constructed. These feature specific kernels help to identify hotspots with more accuracy [9]. MCMI is Information Theory based technique which extracts features with high information and takes care that redundancy is less, making the process fast and efficient [4]. Encoder-Decoder feature extractor consists of convolution and deconvolution layers, which helps features to transform and makes it easier to work with CNNs [18]. Feature extraction methods lose the relations among structures, to solve this issue MCCS is used which stores information in form of matrix [13].

Lithography Hotspot Mitigation is a process of reducing risk of hotspots by taking some preventive measures before lithography process such as during Placement or Routing [28, 32-34]. Adjacent patterns may interfere during placement process leading to hotspots, which can be avoided by using multiple patterning [28]. Lithography simulation and Edge Placement Error (EPE) guided routing [33] help in optimizing the layout after routing process and reduce hotspots by a significant amount. EPE map compares the edge shapes of the layout to edge shapes that are intended in form of a matrix and then finds hotspots.

Machine learning techniques used for hotspot detection purpose are SVM [6, 9, 14, 26, 39] Boosting [6, 37], PCA [6, 8], clustering [28, 35, 38, 41], Naive Bayes [35], Bilinear Classifier (Combination of SVM and Ada-boost) [13]. Bilinear classifier is used with MCCS feature extractor in order to preserve the topological relations [13]. Semisupervised techniques have also identified hotspots with high accuracy [24]. Although ML based detectors overcome the weakness of Pattern Matching, but false alarms remain a problem , which leads to exhaustive and costly post-processing [13, 39].

These days Deep Learning based techniques involving CNNs [12, 15], ANN architecture combined with CNNs [4], GANs [20], CNNs with DBSCAN clustering [22], feature extraction followed by CNNs [18, 23] etc. are being used to reduce False Alarms. Following fig. shows ANN architecture combined with CNNs [4]. The basic block consists of 3 interconnected CNN layers and a Max-pooling layer.

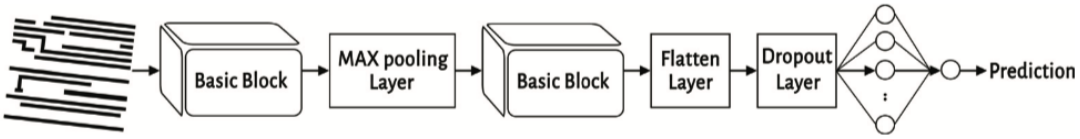


Fig. 2.6 ANN with CNN [4]

Table 2.1 Lithography hotspot detection methods

Category	Sub-Categories	Characterstics	Advantages	Disadvantages
	SAMPLE	• First technique for lithography simulation. • Only for Optical Lithography • Not Available for Personal Computer.		Expensive in terms of

Category	Sub-Categories	Characterstics	Advantages	Disadvantages
Simulation	PROLITH	• Only for Optical Lithography • Available for Personal Computer.	Highly Efficient	Computational time and complexity
	ProMAX, Monte Carlo, ProBEAM	• Available for e beam Lithography		
Pattern Matching	Graph Based	• Large number of False Alarms	Faster than Simulations	Fail to detect previously unseen hotspots
	Template Matching	• High Accuracy • Time taken is high		
	String Search	• Efficient		
Lithography Hotspot Mitigation	Lithography Aware Routing and Placement	• Makes use of EPE and Hotspot maps	Reduces risk of hotspots very significantly	Difficult to perform before the whole process has completed

Category	Sub-Categories	Characterstics	Advantages	Disadvantages
Machine Learning	SVM, Boosting, PCA, Clustering, Semi-supervised technique, Bilinear Classifier, Naive Bayes	• Feature extraction techniques like DBLF, HOLP etc. are used before classifying using these techniques	Detects previously unseen hotspots. Accuracy is high.	False Alarms remain a problem
Deep Learning	CNNs, GANs, CNNs with DBSCAN, CNNs with ANNs	• Most of the techniques implemented till now make use of CNNs	High accuracy and low False Alarms	Computationally expensive

2.2 Based on evaluation parameters:

Based on evaluation parameters there are two basic methods, accuracy matrix and ROC curve.

In accuracy based evaluation two important parameters are evaluated: accuracy and false positives. Accuracy helps us to determine how many hotspots are correctly identified as hotspots and false alarm help us to determine number of non-hotspots determined as hotspots. Ideally, a high accuracy and low false alarms are desired. Mathematically, these can be represented as:

$$Accuracy = \frac{Total\ Hits}{Number\ of\ actual\ hotspots} \tag{2.1}$$

$$FalseAlarm = \frac{Total\ Extras}{Number\ of\ actual\ hotspots} \tag{2.2}$$

Where hits can be defined as a hotspot that has been correctly identified and extra is defined as a non-hotspot that has mistakenly been classified as hotspot [9].

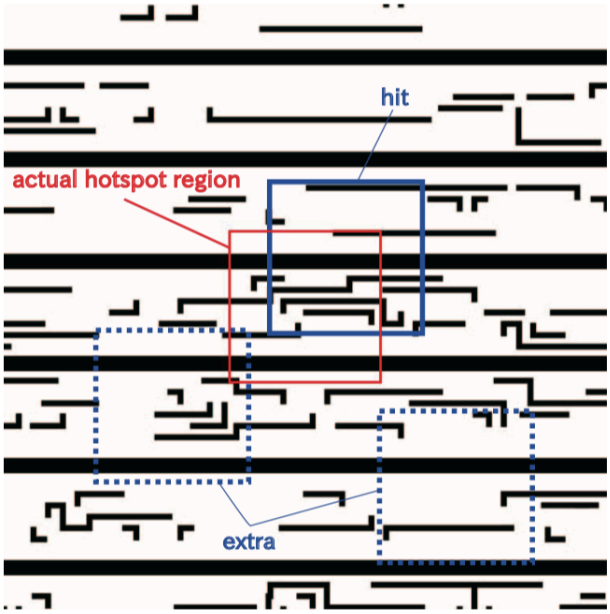


Fig. 2.7 Hit and Extra [7]

After passing the layout through RETs, the number of hotspots left are quiet few in number, however these few can cause high damage to our process hence, it is necessary to detect these. Since number of non-hotspots is much larger than the number of hotspots, the dataset is highly imbalanced. A new technique based on Receiver Operating Characteristics (ROC) can be used to handle this dataset imbalance problem [5].

For this process, an ROC curve is drawn. This curve should show the relationship between the rate of both true positives and true negatives. From this graph AUC and partial AUC scores are obtained. AUC determines number of times the True Positive Rate (TPR) has higher rank than False Positive Rate (FPR), hence comparing accuracy over entire range. Partial AUC takes care of the false alarms. Classifier with higher will have higher AUC score and should be considered better, but as shown in following fig., sometimes that classifier may have lesser FPR in a particular region, leading to false positives. Hence AUC score in that area or particular area of interest is calculated rather than under complete curve, this is known as partial AUC score. For defining these terms mathematically, we can write:

$$\text{AUC Score} = \text{Area under ROC} = \int ROC \tag{2.3}$$

$$\text{Partial AUC Score} = \text{Area under ROC from a to b} = \int_a^b ROC \tag{2.4}$$

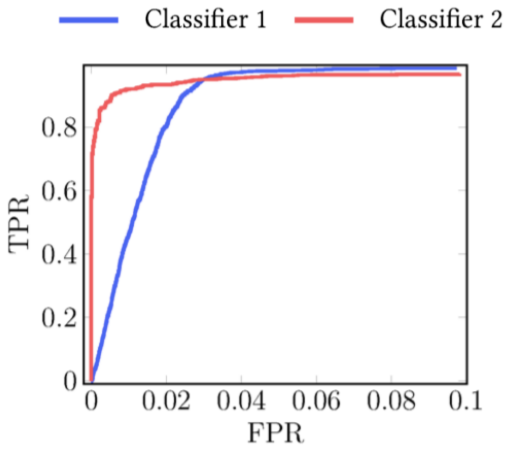


Fig. 2.8 ROC curve for 2 classifiers

This method showed better results in terms of false alarms as it is spread over whole distribution [5].

CHAPTER 3

OVERVIEW

3.1 Lithography

Litho means sculpture of a 2-D or 3-D structure made of metals or stones. Reprinting what we see somewhere else is known as lithography. In IC fabrication process one of the most important steps is lithography a.k.a patterning, which is the process of reprinting the pattern of mask on Silicon wafer. In 1960s circuits were in micrometer range, today these are few nanometers and it has been possible with the help of patterning. It controls shapes dimensions and placement of various components.

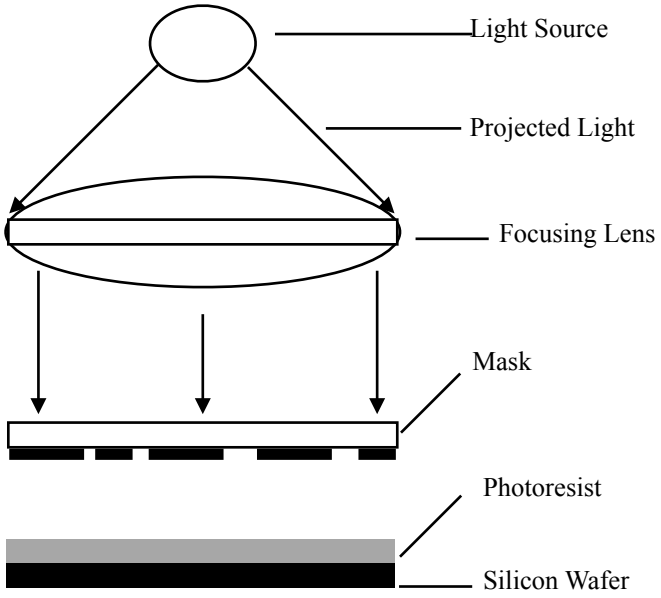


Fig. 3.1 Optical Lithography Process

Based on the type of radiation, there are basically four types of lithography processes:

- Optical Lithography
- Electron beam Lithography
- X-Ray Lithography

- Ion Beam Lithography

Optical lithography, similar to photographic printing, it creates patterns in a layer of photoresist that coats a prepared silicon wafer. Light from an illuminator is projected through a mask that contains the pattern to be created on the wafer. The light patterns. that pass through the mask are reduced by a factor of four by a focusing lens and projected onto the photoresist-coated wafer. Process upto this part is temporary and is known as developing and is reversible. The photoresist part that is exposed to the light becomes soluble or insoluble depending on the type and soluble part is rinsed away, leaving a miniature image of the required pattern at each chip location. This part is permanent, hence needs to be done very carefully.

Mask a.k.a. Reticle contains hardcopy of the design that needs to be transferred from mask to photoresist. For multilayer design each layer has its own mask. Initially, a composite mask is prepared, and later it is broken up in individual mask drawings. It is difficult to design complex masks, so many CAD tools are used for this purpose. For preparing mask, a coating is applied on the glass and pattern is written on it from digital copy using laser. If feature size is greater than 5 microns emulsion coated glass plate is used, if it is less than 5 microns chrome coated glass plate is used. The process of making original hard mask is complex and may take hours to make, but once it is done the pattern can be quickly transferred to different wafers, making it easy to get multiple patterns [43].

Photoresist is the light sensitive material which is applied on top of Silicon wafer. For this, we take one or two drops of photoresist, put it on wafer and spin the wafer really fast so that it uniformly spreads. If the exposed part of resist gets softened and becomes soluble, it is known as positive photoresist and if it becomes insoluble photoresist is negative type. It is made up of two parts a polymer and a photo-sensitive compound. This compound gets activated on exposure to radiation, it absorbs that energy and transfers it to polymer. In negative photoresist, the polymer after getting energy from photoresist, promotes cross-linking leading to increase in molecular weight and decreased solubility. Cross-linking does not take place in presence of oxygen, so nitrogen is used. Also, in negative photoresist, the unexposed regions swell which leads to comparatively poor resolution. In positive photoresist, the molecules of polymer break when exposed to light. Hence, the exposed part

becomes soluble. To make sure part is properly dissolved, exposure time is high.So, in positive photoresist resolution is better, but throughput is high.

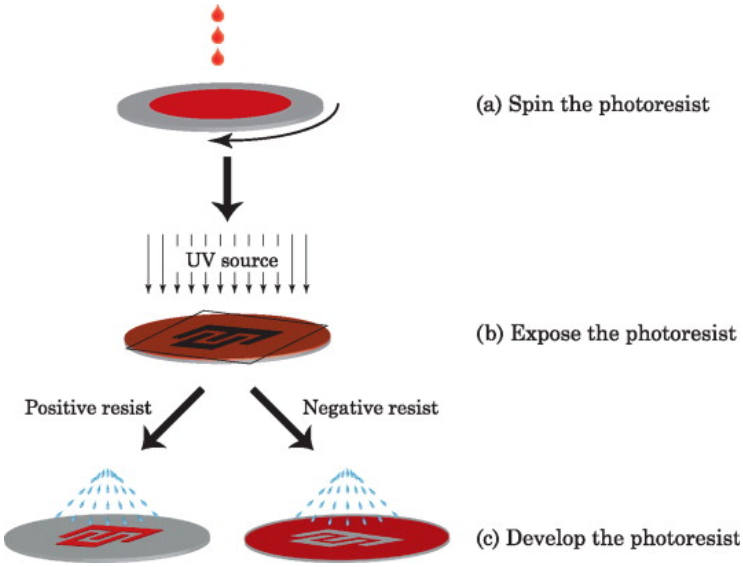


Fig. 3.2 Photoresist

This photoresist is now dipped in developer solution to remove softened part, then etching takes place to remove oxide layer, which prepares final wafer.

Optical Lithography can further be divided into three types: Contact printing, Proximity printing and Projection printing. In contact printing wafer is in contact with mask, hence the resolution is high. But life of the mask is reduced due to wear and tear because of contact. In this type risk of contamination is also there, because if some dirt is present on mask, it is transferred on wafer. In Proximity printing mask and wafer are close, but not in contact. In this type resolution decreases little bit but life increases. In Projection printing, mask and wafer can be as far as we want. In this, a highly focussed image of the mask is projected on wafer, hence resolution is high and wafer life increases. Because of extra optical setup required, cost also increases.

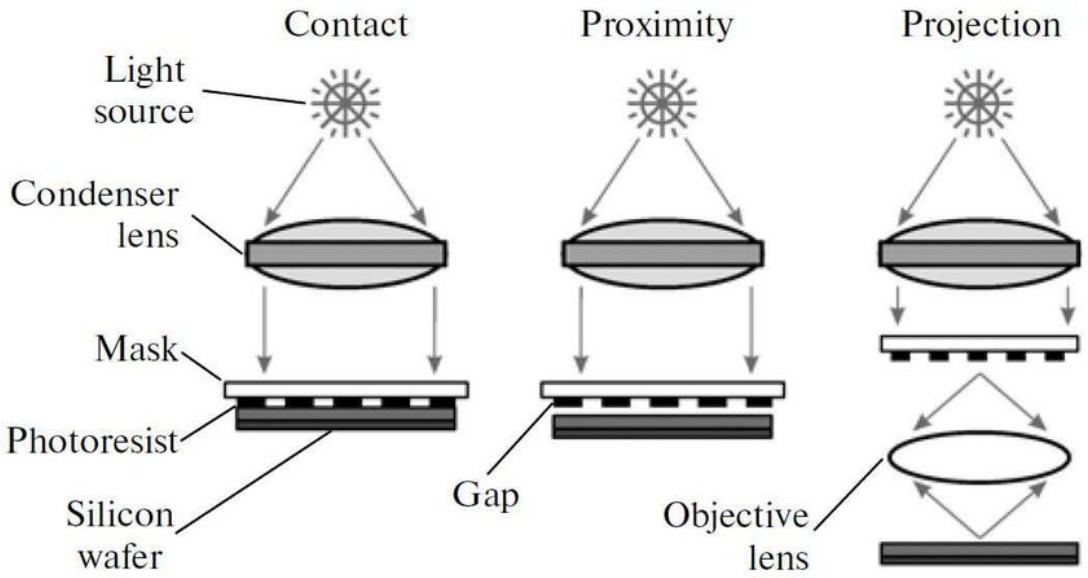
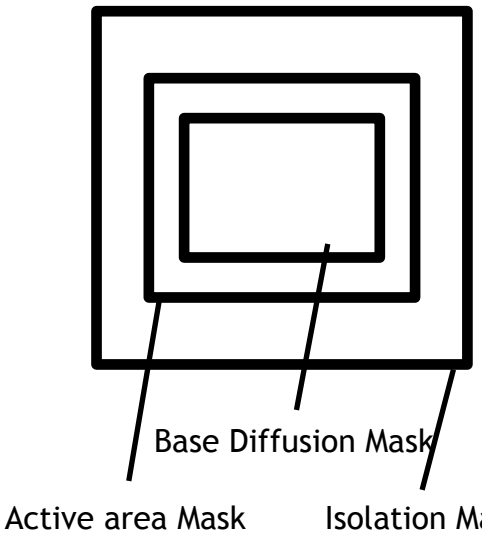


Fig. 3.3 Types of Optical Lithography [42]

There are three figures of merit of lithography:

- Resolution
- Throughput
- Depth of focus

If there is a defined feature size in mask, after transferring it may increase or decrease. The closer the feature size to mask, better the resolution. If original feature size is small, it is required to be very accurate for good resolution. Throughput defines the number of wafers that can be prepared in a given time. It is important because process needs to be cost effective. Since, there are multiple patterns in a design, all of these should be properly aligned with respect to each other as well as base Silicon wafer. For, this alignment to become easier, depth of focus must be good. e.g.: In a simple BJT five masks are required: Active area mask, Junction Isolation mask, Base diffusion mask, Emitter diffusion mask, Contact metal mask. All these masks must be aligned to each other [43].



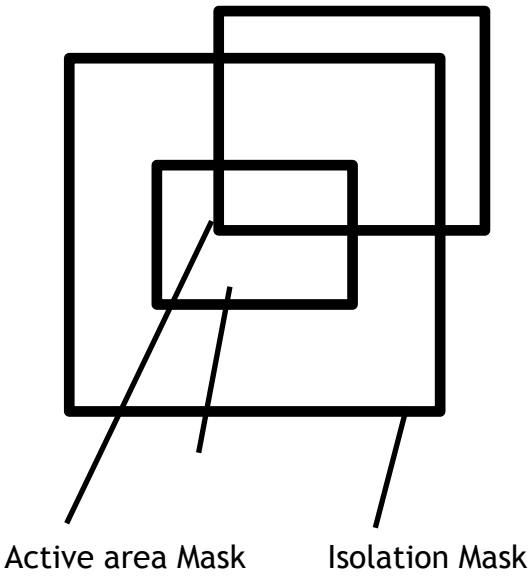


Fig. 3.4 a). Correctly aligned mask

Fig. 3.4 b). Incorrectly aligned mask

Lithography enables Moore's law to be satisfied, which means year by year chip area should go down by half and information we wish to store is increasing. For this to happen, transistors need to get smaller and we need to understand how resistance and capacitance of the wire changes with size. Also 40 to 50 percent of total cost of making chip goes to lithography, so, we wish to continue trend of area getting smaller in a cost-effective way. Angle of diffraction from mask is sinθ

$$sin\theta = \frac{c \,.\, \lambda}{f} \qquad (3.1)$$

where, sinθ is the angle of diffraction,

f is the feature size,

c is constant

λ is the wave length,

0th differential order does not contain any information, so, lens needs to capture at-least first differential order to recreate pattern of the mask. Hence, smaller the feature size, bigger the diffraction angle, bigger the lens we need. As, the feature size decreases lens needs to get bigger and bigger.

Feature size can be defined mathematically as:

$$f = \frac{C \cdot \lambda}{n} \tag{3.2}$$

Here, f is the feature size,

C is Rayleigh constant,

λ is the wave length,

and n is the numerical aperture.

To reduce f, we can increase n or reduce λ.

To increase n, water or oil can be used as medium instead of air. It is known as immersion lithography.

To reduce λ instead of UV light used in optical lithography, we use X-rays. It is known as X-ray lithography and is done using proximity printing. This technique has various advantages such as smaller throughput, low proximity effect and high resolution. X-rays don't absorb dirt so the risk of contamination is also less. When electron beam is focussed on water cooled palladium target, X-rays are generated. This whole process is done inside Helium chamber as, it doesn't absorb X-rays. Mask used in this technique is a thin membrane made of Aluminium Oxide or Silicon or Silicon Nitride, coated with gold (because gold can absorb X-rays). Most commonly used mask for this lithography is Poly Methyl Meta Crystal (PMMA), electron beam resists can also be used. One limitation of this process is blurring of image on the substrate and it depends on the distance between X-ray source and mask and separation of mask and wafer [45].

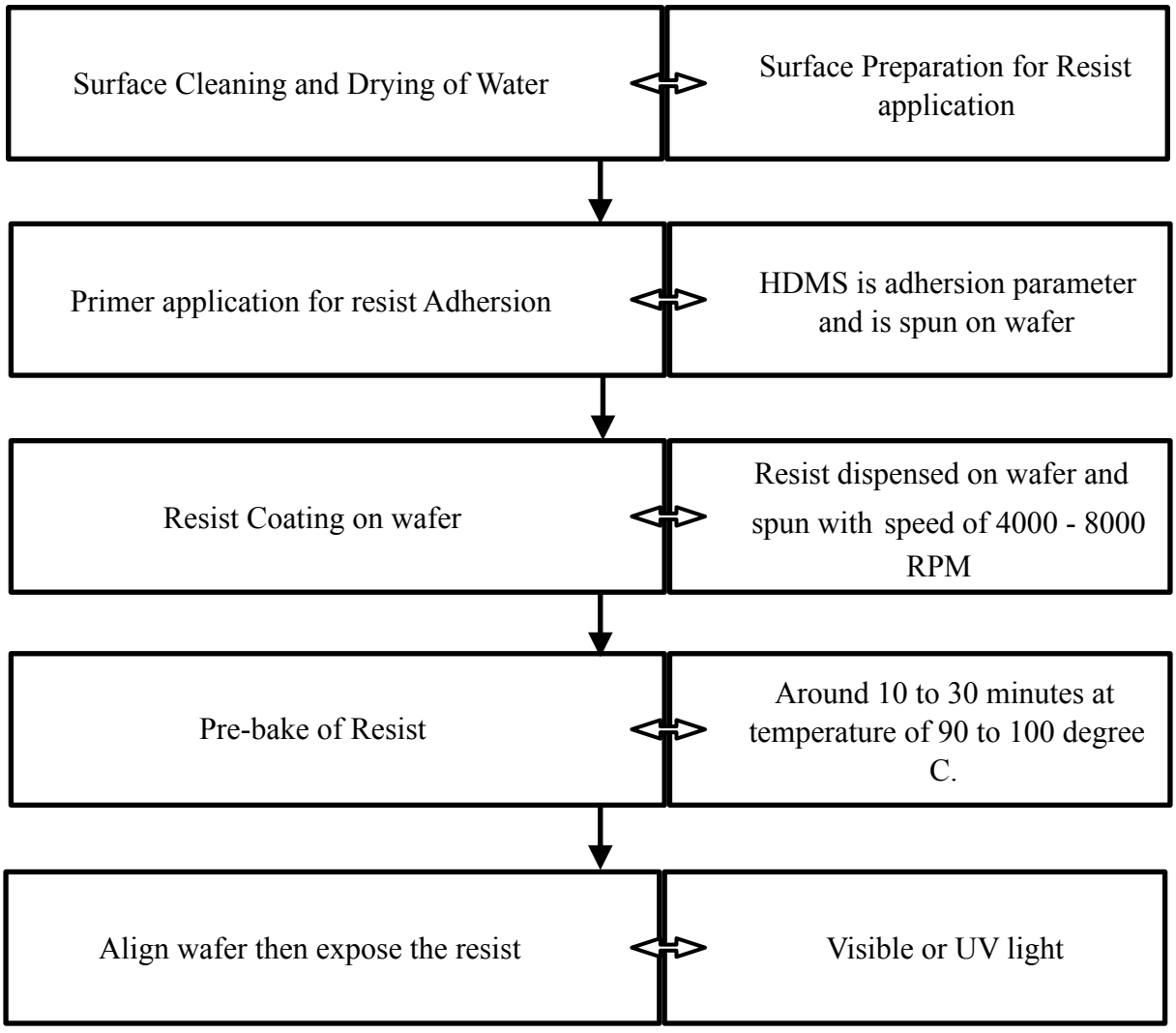
Another way of obtaining smaller feature size is by using electron beam lithography. In this direct writing on the substrate is possible and mask is not required. Hence, it provides better depth of focus, resolution and can be easily automated. To perform this technique e^- beam is focussed on substrate and focussed beam is scanned on some area of substrate. That area is defined as scan field. Beam is turned on or off depending on where pattern is required or not. After one scan field is completed, substrate moves. Beam diameter should be small and focussed. Ideally minimum feature size is four times the beam diameter and scan field is 200 times the beam diameter. The time required for this is very high, so it leads to a smaller throughput. Another disadvantage of this lithography technique is proximity effect. When electron beam is focussed on substrate, scattering of electrons take place so, they go large distances away from original pattern. Because of this broadening of actual pattern and feature

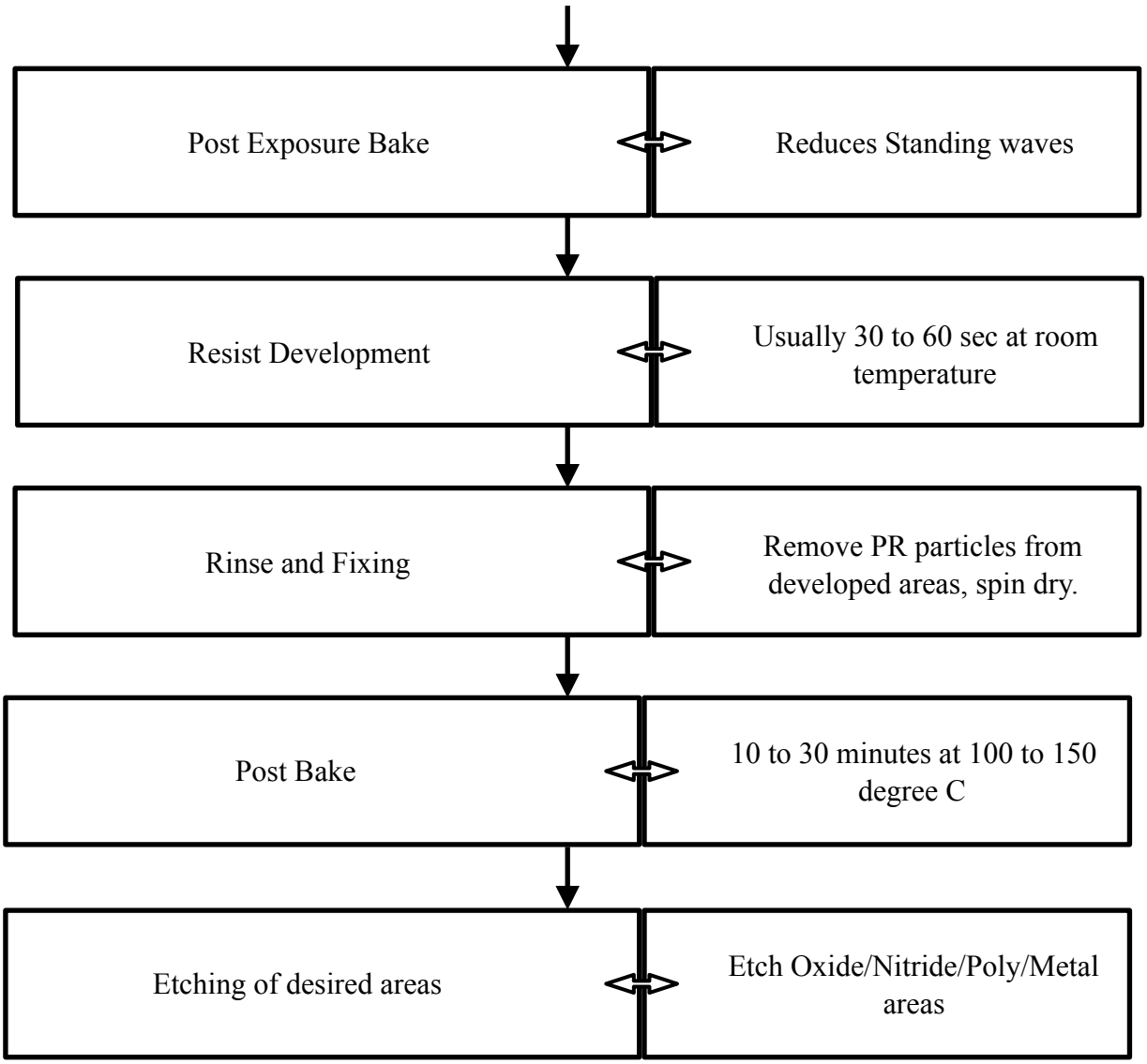
size takes place, this is known as proximity effect. Insufficient radiation at corners doesn't let them to fully develop, this effect is known as intra-proximity and over development of certain areas due to extra exposure is called inter-proximity [43].

Ion-beam can be used in place of electron beam which leads to less scattering and hence, less proximity effect. This process is known as Ion-beam lithography.

Designers are given certain rules such as how much edge to edge distance one should keep, how much separation should be there, length and width etc. which must be followed to get least errors possible. But sometimes because of decreasing feature size printability issues occur, those places with such issues are known as lithography hotspots.

Table 3.1 Optical Lithography Process [44]





3.2 Vision Transformer

Today most of the best performing Natural Language Processing models are Transformer based, but in computer vision based applications, these are not used to much extent. Till now transformers have been used with CNNs, never at their place. Vision Transformers aim to use transformers for image classification tasks without involvement of convolutions [3].

Ideally, transformers operate on sequences or sets. Transformer applies attention mechanism on the sets. Since attention is a quadratic operation, we have to calculate pair-wise inner product between each pair of sets, therefore computations and memory required are very high. Images are harder because these are composed of many pixels. Even a small image consists of 250*250 pixels. Every pixel needs to attend every other pixel, so even for a small

image we will need $((250)^2)^2$ operations. This much computations are not possible to achieve with hardware [2].

Vision transformers make use of transformers by including some extra operations, which are explained further.

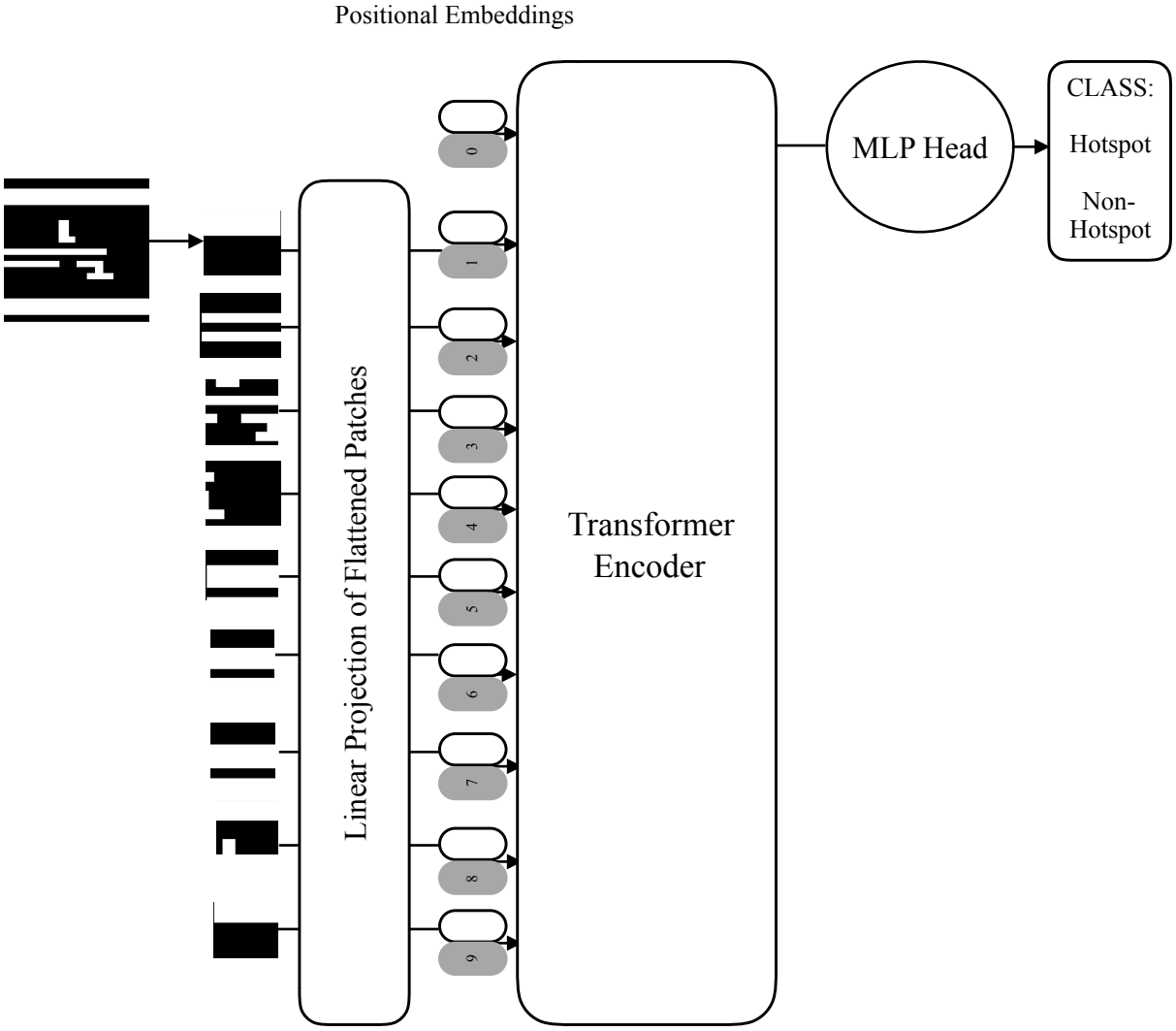


Fig. 3.5 Structure of ViT [3]

First of all, image is partitioned into patches of same shape. These patches may or may not overlap, every patch is a small image. Then these patches are vectorized by reshaping tensors to vectors, let the output be $X_1, X_2, \ldots, X_n$. Next dense layer is applied to these vectors, let the outputs be $Z_1, Z_2, \ldots, Z_n$, where $Z_n = w*X_n + b$. Then, positional encoding is added to these patches, because swapping of patches may lead to information loss. Other than these CLS token is passed through embedded layer and its output is used to provide classification

output. All these vectors are passed through Transformer encoder network. Its first output is feature vector, which is passed through MLP head, which acts as a classifier and provides image classification output.

In transformers we can pay attention to the things which are far away even at beginning from end, 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 we integrate over a pixel, which connects to its neighbourhood. Then that neighbourhood connects to its neighbourhood 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 [2].

CHAPTER 4

EXPERIMENTS PERFORMED

4.1 Dataset Used

For this research, we have utilized ICCAD-2012 dataset. It has five sub data-sets with different types of layouts. First dataset is obtained using 32 nm process and other four have been obtained during 28 nm process. Each sub-dataset contains training and test set, detailed information of each sub-dataset is provided in the following table.

Table 4.1 Details of the dataset used

	TRAINING SET			TEST SET		
	Hotspot	Non-Hotspot	Total	Hotspot	Non-Hotspot	Total
Sub-dataset 1	99	340	439	226	3869	4095
Sub-dataset 2	174	5285	5459	498	41298	41796
Sub-dataset 3	909	4643	5552	1808	46333	48141
Sub-dataset 4	95	4452	4547	177	31890	32067
Sub-dataset 5	26	2176	2202	41	19327	19368
TOTAL	1303	16896	18199	2750	142717	145467

We can see from table 4.1 that total number of images with hotspot in training set and test set are 1303 and 2750 respectively and number of Non-Hotspot images in training and test set are 16896 and 142717 respectively. The results provided are after passing our output through design will check and hence number of hotspots is much less than number of non-hotspots. The data is it is highly imbalanced many techniques have been used to take care of this imbalance, some of them being data augmentation and filtering [1].

Other problems faced by this dataset is false alarms. It has been seen that using synthetic patterns to increase the amount of training data significantly reduces these false alarms [19].

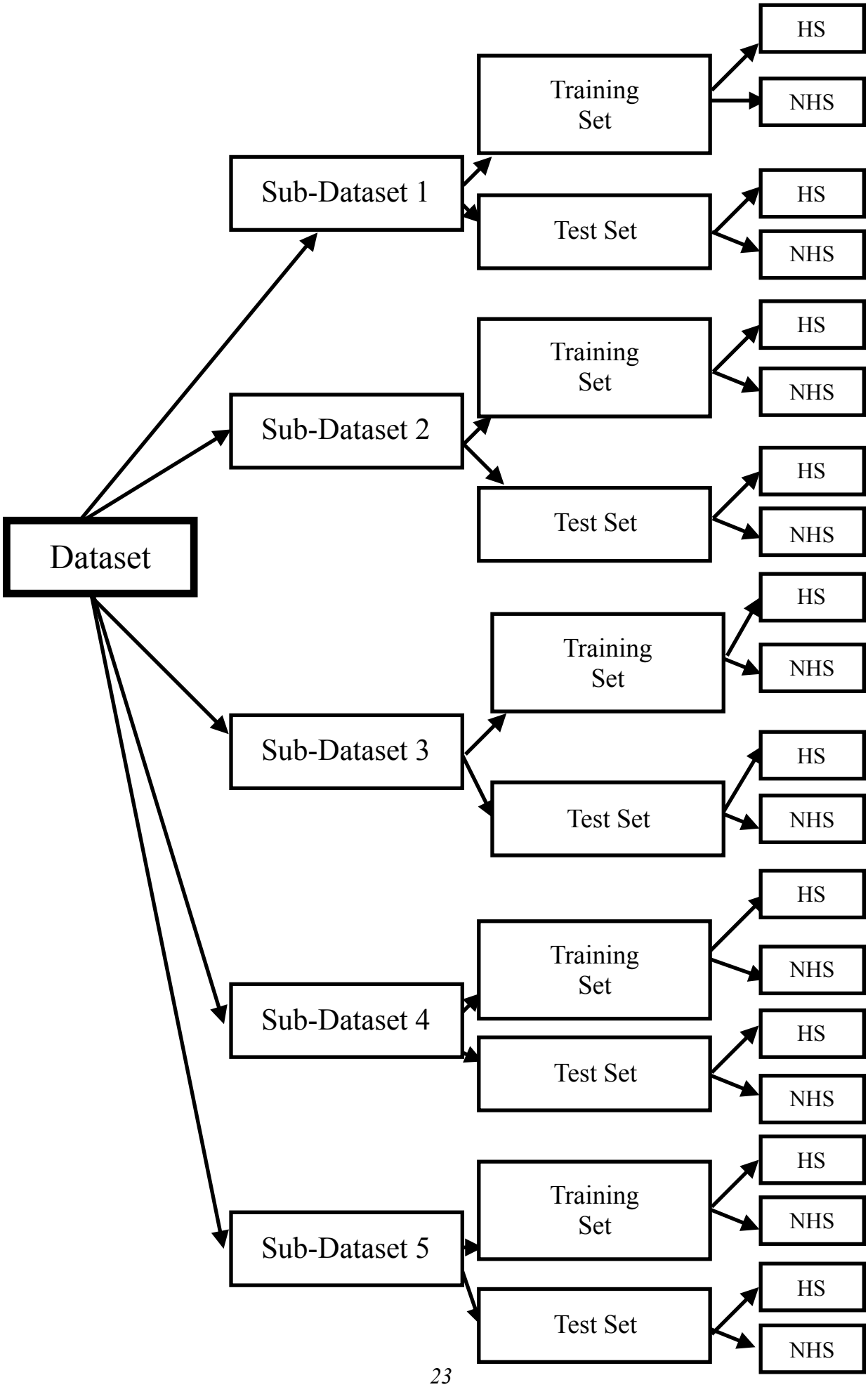

Dataset
Sub-Dataset 1
Sub-Dataset 2
Sub-Dataset 3
Sub-Dataset 4
Sub-Dataset 5
Training Set
Test Set
HS
NHS

Fig. 4.1 Representation of dataset

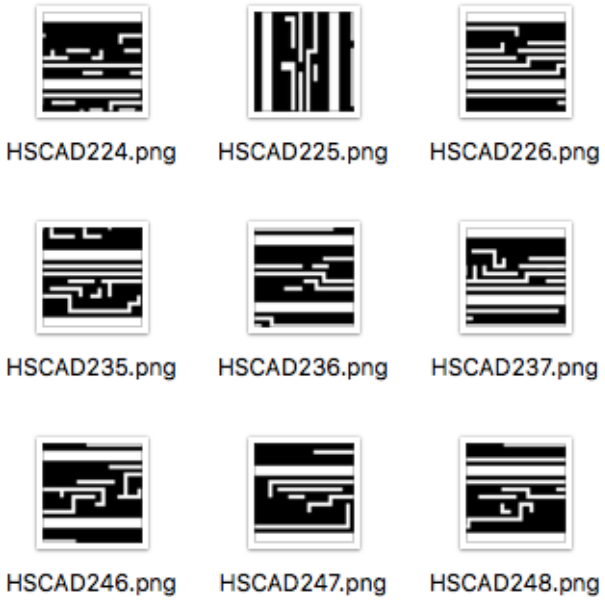


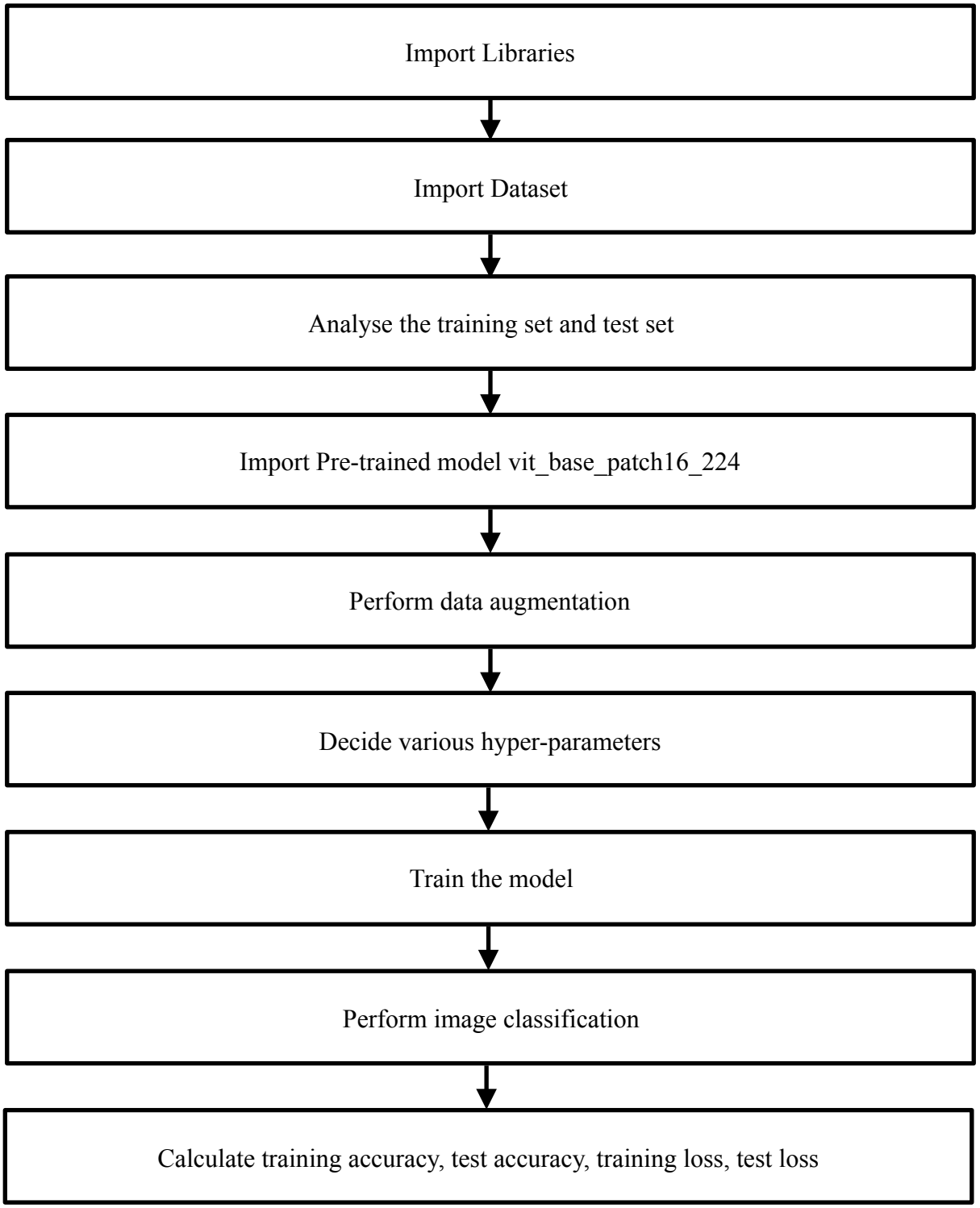
Fig. 4.2 Images with Hotspot

Fig. 4.3 Images without Hotspot

4.2 Model

We aim to detect hotspots using ViT, the specifications of model used is as follows:

Table 4.2 Steps performed for classification using ViT



To implement ViT, we loaded a pre-trained transformer model 'vit_base_patch16_224' provided by Hugging Face. This model is trained on ImageNet21k dataset and then fine-tuned on ImageNet dataset. It uses a patch size of 16*16. Chained transformations are used to resize images to 224*224 resolution. These resized images are then converted to tensors and these tensors are normalized with value of mean and standard deviation = 0.5 for all three channels. Model has been created using torch image model library and pre-trained transformer model. For generating the model batch size set to a value of 100, learning rate with values 1e-2 for datasets 2,3,4,5 and 1e-3 for dataset1 gave the best results. While compiling the model, Adam optimizer, cross-entropy loss and accuracy matrix have been used and the output of the model is discussed in results section.

ViT hyper-parameters used:

- Batch_size = 100
- Number of epochs = 15
- Image Size = 224
- Learning Rate: 1e-3 for sub-dataset 1, 1e-2 for sub-dataset 2,3 and 4, 5e-2 for sub-dataset 5
- Epochs < 5
- Patch size = 16x16

For comparison purposes we performed the experiment using CNNs and ANNs as well and used Tensor flow Keras library. For generating model for CNNs and ANNs Sequential model with a batch size of 64 and Images reshaped to 224*224 resolution have been used. For generating model for CNNs 3 Convolution layers with 12 filters and Kernel Size = (3,3) , 2 max-pooling layers with pooling window (2,2) and 2 dense layers give the best results. For ANNs best accuracy has been obtained using 9 dense layers and 8 dropout layers with with rate = 0.3. Same parameters have been used while compiling the models for all three techniques. Steps performed are same as in case of ViT except that we don't use pre-trained model for CNNs.

CNN hyper-parameters used is as follows:

- Model = Sequential
- Batch Size = 64
- Image Shape = 224
- Number of Convolution layers = 3
    - Number of filters = 12
    - Kernel Size = (3,3)
    - Activation = relu
- Number of Max-Pooling layers = 2
    - Pooling window = (2,2)
- Number of dense layers = 2
- Number Of Epochs <= 10
- Optmizer = Adam
- Loss = Sparse Categorical Cross-entropy
- Metrics = Accuracy

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 222, 222, 12)      336
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 111, 111, 12)      0
_________________________________________________________________
conv2d_1 (Conv2D)            (None, 109, 109, 12)      1308
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 54, 54, 12)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 52, 52, 12)        1308
_________________________________________________________________
flatten (Flatten)            (None, 32448)             0
_________________________________________________________________
dense (Dense)                (None, 64)                2076736
_________________________________________________________________
dense_1 (Dense)              (None, 10)                650
=================================================================
Total params: 2,080,338
Trainable params: 2,080,338
Non-trainable params: 0
_________________________________________________________________
```

Fig. 4.4 Summary of CNN model

ANN hyper-parameters used is as follows:

- Model = Sequential
- Batch Size = 64
- Image Shape = 224
- Number of Dense layers = 9
    - Activation = relu for first 8 layers and softmax for last layer
- Number of Dropout layers = 7
    - Dropout rate = 0.7
- Number Of Epochs <= 10
- Optmizer = Adam
- Loss = Sparse Categorical Cross-entropy
- Metrics = Accuracy

Layer (type)	Output Shape	Param #
flatten (Flatten)	(None, 150528)	0
dense (Dense)	(None, 2048)	308283392
dropout (Dropout)	(None, 2048)	0
dense_1 (Dense)	(None, 1024)	2098176
dropout_1 (Dropout)	(None, 1024)	0
dense_2 (Dense)	(None, 512)	524800
dropout_2 (Dropout)	(None, 512)	0
dense_3 (Dense)	(None, 256)	131328
dropout_3 (Dropout)	(None, 256)	0
dense_4 (Dense)	(None, 128)	32896
dropout_4 (Dropout)	(None, 128)	0
dense_5 (Dense)	(None, 56)	7224
dropout_5 (Dropout)	(None, 56)	0
dense_6 (Dense)	(None, 28)	1596
dropout_6 (Dropout)	(None, 28)	0
dense_7 (Dense)	(None, 14)	406
dropout_7 (Dropout)	(None, 14)	0
dense_8 (Dense)	(None, 10)	150

Total params: 311,079,968

Fig. 4.5 Summary of ANN model

CHAPTER 5

RESULTS

Table 5.1 Accuracy for all subsets using ViT, CNNs and ANNs

Model	Accuracy					Overall Average Accuracy
	Sub-Dataset 1	Sub-Dataset 2	Sub-Dataset 3	Sub-Dataset 4	Sub-Dataset 5	
ViT	95.48	99.37	95.77	99.83	99.80	98.05
CNN	94.37	98.81	90.91	99.45	99.79	96.666
ANN	89.58	97.73	94.58	98.68	99.48	96.01

From Table 5.1, Comparing in terms of overall accuracy, we can see that for ViT performs better than CNNs and ANNs.

For Sub-datasets 1, 2, 3, 4 and 5 it gives the best results. CNNs perform moderately well for sub-datasets 1, 2, 4 and 5 and worst for sub-dataset 3. ANNs perform poorest for sub-datasets 1, 2, 4, 5 and moderately for sub-dataset 3.

In terms of overall accuracy ViTs give 1.39% better accuracy than CNNs and 2.04% better accuracy than ANNs.

Table 5.2 Comparison with other works

Model	Accuracy					Overall Average Accuracy
	Sub-Dataset 1	Sub-Dataset 2	Sub-Dataset 3	Sub-Dataset 4	Sub-Dataset 5	
Ours (ViT)	95.48	99.37	95.77	99.83	99.80	98.05
Y.Yu et.al, [9]	93.81	98.2	91.88	85.94	92.86	92.538

Model	Accuracy					Overall Average Accuracy
	Sub-Dataset 1	Sub-Dataset 2	Sub-Dataset 3	Sub-Dataset 4	Sub-Dataset 5	
H.Yang et.al, [15]	99.6	99.8	97.8	96.4	95.1	97.74
H.Zhang et.al, [35]	100	99.4	97.52	97.74	95.12	97.956
T. Matsunaw a et. al [37]	100	98.6	97.2	87.1	92.68	95.116

Table 5.2 shows comparison of our ViT model with other researches. In terms of overall accuracy ViT model gives the best result. In terms of sub-datasets, for sub-dataset 4 and 5, ViT gives the best accuracy. For sub-dataset 2, accuracy is not the best of all, but it is comparable to best performing models. For sub-dataset 1, [35] and [37] provide the best results and for sub-dataset 3, [15] gives the best accuracy.

CHAPTER 6

CONCLUSION & FUTURE SCOPE

6.1 Conclusion

In this work we have detected lithography hotspots 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 as well. From table 5.1, we can see that in terms of overall accuracy ViT give 1.39% better accuracy than CNNs and 2.04% better accuracy than ANNs. Considering sub-datasets wise accuracies, ViT performs better or as good as CNNs for each sub-dataset. For three out of five sub-datasets accuracy on the test set is more than 99% and for other two it is more than 95%. We also compared our work to already existing works and it can be see from table 5.2 that in terms of overall accuracy, ViT gives the best results and at sub-dataset level for three out of five, it provides best or comparable results but lags for two sub-datasets. From the results we can conclude that although the proposed technique performs better than a lot of already existing state of the art techniques, it fails to supplant all the existing methods for all the sub-datasets. ViTs can be seen as an new and alternate method for the purpose of identifying hotspots in lithography. Since the technique is very novel, many improvements lie for it in the coming future. With the improvements in technique, increase in accuracy and lower time can also be expected for our problem statement.

6.2 Future Scope

Presence of hotspots can lead to complete failure of the circuit. In order to avoid this we wish to obtain as high accuracy as possible. Our aim for future researches remains to reduce the iteration time in this technique and improve accuracy for sub-datasets 1 and 3 by improving our model and modifying the dataset using techniques like mirror flipping, upsampling etc. to reduce the imbalance in it and at the same time increasing training data.

REFERENCES

[1]. J. A. Torres, "ICCAD-2012 CAD contest in fuzzy pattern matching for physical verification and benchmark suite," 2012 IEEE/ACM International Conference on Computer-Aided Design (ICCAD), 2012, pp. 349-350.

[2]. S. Gkelios, Y. Boutalis, S.A. Chatzichristofis “Investigating the Vision Transformer Model for Image Retrieval Tasks,” 11 Jan 2021, [Online], Available: arXiv:2101.03771v1. [Accessed: June 2021]

[3]. A. Dosovitskiy, L. Beyer, A. Kolesnikov, D.Weissenborn, X. Zhai, T.Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, N.Houlsby “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” 3 Jun 2021, [Online], Available: arXiv:2010.11929v2. [Accessed: June 2021]

[4]. V. Borisov and J. Scheible, "Lithography Hotspots Detection Using Deep Learning," 2018 15th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), 2018, pp. 145-148, doi: 10.1109/SMACD.2018.8434561.

[5]. W. Ye , Y. Lin , M.Li , Q. Liu , D.Z. Pan, “LithoROC: lithography hotspot detection with explicit ROC optimization,” ASPDAC '19: Proceedings of the 24th Asia and South Pacific Design Automation Conference, January 2019, pp. 292–298, doi.: 10.1145/3287624/3288746.

[6]. B. Yu, J.R. Gao, D. Ding, X. Zeng, D. Z. Pan, "Accurate lithography hotspot detection based on principal component analysis-support vector machine classifier with hierarchical data clustering," J. Micro/Nanolith. MEMS MOEMS, 4 November 2014, doi.: 10.1117/1.JMM.14.1.011003.

[7]. Y. Tomioka, T. Matsunawa, C. Kodama and S. Nojima, "Lithography hotspot detection by two-stage cascade classifier using histogram of oriented light propagation," 2017

22nd Asia and South Pacific Design Automation Conference (ASP-DAC), 2017, pp. 81-86, doi: 10.1109/ASPDAC.2017.7858300.

[8]. W. Wen, J. Li, S. Lin, J. Chen and S. Chang, "A Fuzzy-Matching Model With Grid Reduction for Lithography Hotspot Detection," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 33, no. 11, pp. 1671-1680, Nov. 2014, doi: 10.1109/TCAD.2014.2351273.

[9]. Y. Yu, G. Lin, I. H. Jiang and C. Chiang, "Machine-Learning-Based Hotspot Detection Using Topological Classification and Critical Feature Extraction," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 34, no. 3, pp. 460-470, March 2015, doi: 10.1109/TCAD.2014.2387858.

[10]. G. Huang, J. Hu, Y. He, J. Liu, M.Ma, Z. Shen, J. Wu, Y. Xu, H. Zhang, K. Zhong, X. Ning, Y. Ma, H. Yang, B.Yu, H. Yang, Y. Wang “Machine Learning for Electronic Design Automation: A Survey,” 8 Mar 2021, [Online], Available: arXiv:2102.03357v2arXiv: 2102.03357v2. [Accessed: June 2021]

[11]. G. R. Reddy, C. Xanthopoulos and Y. Makris, "Enhanced hotspot detection through synthetic pattern generation and design of experiments," 2018 IEEE 36th VLSI Test Symposium (VTS), 2018, pp. 1-6, doi: 10.1109/VTS.2018.8368646.

[12]. Y Xiao., X. Huang, and K. Liu, “Transferability from ImageNet to Lithography Hotspot Detection,” J Electron Test 37, pp. 141–149, doi.: 10.1007/s10836-021-05925-5.

[13]. H. Zhang , F. Zhu , H. Li , E.F.Y. Young, B. Yu “Bilinear Lithography Hotspot Detection Share,” ISPD '17: Proceedings of the 2017 ACM on International Symposium on Physical Design, March 2017, Pp. 7–14, doi.: 10.1145/3036669.3036673.

[14]. H. Yang, Y. Lin, B. Yu, E.F. Y. Young, “Lithography Hotspot Detection: From Shallow To Deep Learning”, IEEE International System-on-Chip Conference (SOCC), pp. 233–238, Munich, Germany, September 5–8, 2017.

[15]. H. Yang, L. Luo, J. Su, C. Lin, B. Yu, “Imbalance Aware Lithography Hotspot Detection: A Deep Learning Approach”, SPIE Intl. Symp. Advanced Lithography Conference, San Jose, CA, Feb. 26–Mar. 2, 2017.

[16]. J. A.T. Robles, S. Mostafa, K. Madkour, J.Y. Wuu, “Hotspot Detection Based on Machine Learning,” United States Patent Application Publication, 2013, , [Online], Available: https://patents.google.com/patent/US20130031522. [Accessed: June 2021]

[17]. Z. XingYu, Y. YouLing “Hotspot Detection of Semiconductor Lithography Circuits Based on Convolutional Neural Network,” Journal of Microelectronic Manufacturing, December 2018, pp. 1-8, doi.: 10.33079/jomm.18010205.

[18]. R. Chen, W. Zhong, H. Yang, H. Geng, X. Zeng and B. Yu, "Faster Region-based Hotspot Detection," 2019 56th ACM/IEEE Design Automation Conference (DAC), 2019, pp. 1-6.

[19]. G. R. Reddy, C. Xanthopoulos and Y. Makris, "On Improving Hotspot Detection Through Synthetic Pattern-Based Database Enhancement," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, doi: 10.1109/TCAD. 2021.3049285.

[20]. W. Ye, M. B. Alawieh, Y. Lin, D.Z. Pan, “LithoGAN: End-to-End Lithography Modeling with Generative Adversarial Networks,” The 56th Annual Design Automation Conference 2019, June 2019, doi.:10.1145/3316781.3317852.

[21]. S. Tamagawa, R. Fujimoto, M. Inagi, S. Nagayama, S. Wakabayashi, “A Hotspot Detection Method Based on Approximate String Search,” CENICS 2016: The Ninth International Conference on Advances in Circuits, Electronics and Micro-electronics, 2016.

[22]. M. Shin, J.H. Lee, "Accurate lithography hotspot detection using deep convolutional neural networks," J. Micro/Nanolith. MEMS MOEMS 15(4), 18 Nov 2016, doi.: 10.1117/1.JMM.15.4.043507.

[23]. H. Yang, J. Su, Y. Zou, Y. Ma, B. Yu and E. F. Y. Young, "Layout Hotspot Detection With Feature Tensor Generation and Deep Biased Learning," in IEEE Transactions on

Computer-Aided Design of Integrated Circuits and Systems, vol. 38, no. 6, pp. 1175-1187, June 2019, doi: 10.1109/TCAD.2018.2837078.

[24]. Y. Chen, Y. Lin, T. Gai, Y. Su, Y. Wei and D. Z. Pan, "Semisupervised Hotspot Detection With Self-Paced Multitask Learning," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 39, no. 7, pp. 1511-1523, July 2020, doi: 10.1109/TCAD.2019.2912948.

[25]. Y. Yu, Y. Chan, S. Sinha, I. H. Jiang and C. Chiang, "Accurate process-hotspot detection using critical design rule extraction," DAC Design Automation Conference 2012, 2012, pp. 1163-1168.

[26]. W. Ye, M. B. Alawieh, M. Li, Y. Lin and D. Z. Pan, "Litho-GPA: Gaussian Process Assurance for Lithography Hotspot Detection," 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE), 2019, pp. 54-59, doi: 10.23919/DATE.2019.8714960.

[27]. K. Liu, B. Tan, R. Karri and S. Garg, "Poisoning the (Data) Well in ML-Based CAD: A Case Study of Hiding Lithographic Hotspots," 2020 Design, Automation & Test in Europe Conference & Exhibition (DATE), 2020, pp. 306-309, doi: 10.23919/DATE48585.2020.9116489.

[28]. J. Gao, B. Yu, D. Ding and D. Z. Pan, "Lithography hotspot detection and mitigation in nanometer VLSI," 2013 IEEE 10th International Conference on ASIC, 2013, pp. 1-4, doi: 10.1109/ASICON.2013.6811917.

[29]. C.A. Mack, "Thirty years of lithography simulation," Optical Microlithography XVIII, 12 May 2005, doi.: 10.1117/12.601590.

[30]. C. Chiang and J. Kawa, "Three DFM Challenges: Random Defects, Thickness Variation, and Printability Variation," APCCAS 2006 - 2006 IEEE Asia Pacific Conference on Circuits and Systems, 2006, pp. 1099-1102, doi: 10.1109/APCCAS.2006.342313.

[31]. J. Xu, S. Sinha and C. C. Chiang, "Accurate detection for process-hotspots with vias and incomplete specification," 2007 IEEE/ACM International Conference on Computer-Aided Design, 2007, pp. 839-846, doi: 10.1109/ICCAD.2007.4397369.

[32]. S.K. Kim, J.E. Lee, S.W. Park, H.K. Oh, “Optical lithography simulation for the whole resist process,” Current Applied Physics, Volume 6, Issue 1, 2006, pp. 48-53, doi.: 10.1016/j.cap.2004.12.003.

[32]. D.Z. Pan, P.Y. Huazhong, M. Cho, A. Ramalingam, “Design for manufacturing meets advanced process control: A survey,” Journal of Process Control, December 2008, doi.: 10.1016/j.jprocont.2008.04.007.

[33]. J. Mitra , P. Yu , D.Z. Pan “RADAR: RET-aware detailed routing using fast lithography simulations,” DAC '05: Proceedings of the 42nd annual Design Automation Conference, June 2005, pp. 369–372 doi.: 10.1145/1065579.1065678.

[34]. D. Z. Pan, "Lithography-aware physical design," 2005 6th International Conference on ASIC, 2005, pp. 1172-1173, doi: 10.1109/ICASIC.2005.1611242.

[35]. H. Zhang , B. Yu , E.F. Y. Young, “Enabling online learning in lithography hotspot detection with information-theoretic feature optimization,” ICCAD '16: Proceedings of the 35th International Conference on Computer-Aided Design, November 2016, pp. 1–8 doi.: 10.1145/2966986.2967032.

[36] S. Tamagawa, R. Fujimoto, M. Inagi, S. Nagayama, S. Wakabayashi, “ATable Reference-Based Acceleration of a Lithography Hotspot Detection Method Based on Approximate String Search,” CENICS 2017: The Tenth International Conference on Advances in Circuits, Electronics and Micro-electronics, 2017.

[37] T. Matsunawa, J.R. Gao, B. Yu, and D. Z. Pan, “A new lithography hotspot detection framework based on AdaBoost classifier and simplified feature extraction,” SPIE vol. 9427, 2015.

[38] N. Ma, J.Ghan, S. Mishra, C. Spanos, K. Poolla, N. Rodriguez, L. Capodieci, “Automatic hotspot classification using pattern-based clustering,” SPIE Design for Manufacturability through Design-Process Integration II, 4 March 2008, doi.: 10.1117/12.772867.

[39] D. Ding, J. A. Torres and D. Z. Pan, "High Performance Lithography Hotspot Detection With Successively Refined Pattern Identifications and Machine Learning," in IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, vol. 30, no. 11, pp. 1621-1634, Nov. 2011, doi: 10.1109/TCAD.2011.2164537.

[40] https://nptel.ac.in/courses/117/106/117106093/

[41] J.Ghan, N.Ma, S. Mishra, C. Spanos, K. Poolla, N. Rodriguez, L. Capodieci, "Clustering and pattern matching for an automatic hotspot classification and detection system," SPIE Design for Manufacturability through Design-Process Integration III, 12 March 2009, doi.: 10.1117/12.814328.

[42] Haberfehlner, Georg “3D nanoimaging of semiconductor devices and materials by electron tomography”,2013.

[43] Prof. Nandita Dasgupta “ Lithography Lecture-1 and Lecture-1”, [Online], Available: https://nptel.ac.in/courses/117/106/117106093/. [Accessed: Jully 2021]

[44] Prof. AN Chandokar “ Lecture 16 to 18, Lithography”, [Online], Available:https:// nptel.ac.in/courses/108/101/108101089/. [Accessed: Jully 2021]

[45] Prof. Parasuraman Swaminathan “ Lithography”, [Online], Available: https:// onlinecourses.nptel.ac.in/noc20_mm25/. [Accessed: Jully 2021]

[46] S. Sivakumar, "EUV lithography: Prospects and challenges," 16th Asia and South Pacific Design Automation Conference (ASP-DAC 2011), 2011, pp. 402-402, doi: 10.1109/ ASPDAC.2011.5722221.

LIST OF PUBLICATIONS

[1]. "Empirical Laws of Natural Language Processing for Neural Language Generated Text,"Springer CCIS 6th International Conference Information, Communication & Computing Technology, Presented on: May 8th 2021.

Springer CCIS ICICCT-2021 jims

Certificate of Presentation

Awarded to

Sumedha

for presenting the paper,

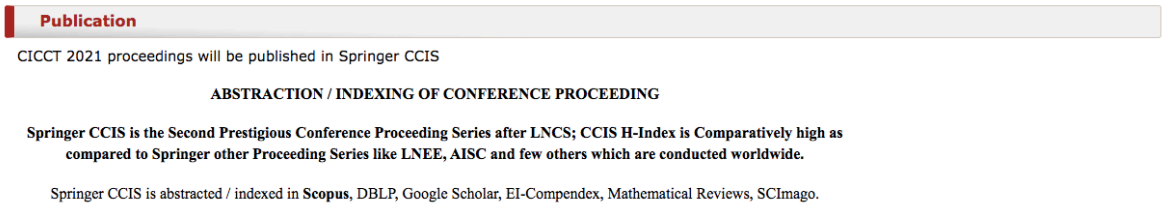
Empirical Laws Of Natural Language Processing For Neural Language Generated Text

at 6th International Conference on Information, Communication & Computing Technology (ICICCT - 2021) in collaboration with Springer CCIS and Computer Society of India, held on 8th May 2021 conducted by Jagan Institute of Management Studies (JIMS) Sector-05, Rohini, New Delhi, India. Due to ongoing Corona pandemic the Conference is conducted in live stream mode through Zoom. Selected & presented papers will be published in CCIS Vol. 1417.

Dr. Praveen Arora
Conference Secretariat

Dr. Latika Kharb
Convener

Dr. Deepak Chahal
Convener



Publication

CICCT 2021 proceedings will be published in Springer CCIS

ABSTRACTION / INDEXING OF CONFERENCE PROCEEDING

Springer CCIS is the Second Prestigious Conference Proceeding Series after LNCS; CCIS H-Index is Comparatively high as compared to Springer other Proceeding Series like LNEE, AISC and few others which are conducted worldwide.

Springer CCIS is abstracted / indexed in **Scopus**, DBLP, Google Scholar, EI-Compendex, Mathematical Reviews, SCImago.

[2]. "Exploring Vision Transformer model for detecting Lithography Hotspots" submitted in Defence Science Journal