機器學習在半導體製程之光學OPC驗證準確度提升之應用

Application of Machine learning (ML) in the Manufacturing Process of Semiconductors-- To improve OPC verification accuracy

國立台灣大學 化學工程學所 博士班三年級 曾麒樺
指導教授: 李克強 教授

研究重點

In the field of lithography, the method of correcting the image distortion is called optical proximity correction (OPC). Now the mainstream OPC technology is model-based OPC, which is characterized by the need to collect a huge amount of wafer feature size to build a model. If a hotspot appears, all previous efforts will be wasted. So hotspot repair and model re-correction are important tasks. In addition, how to speed up the model establishment and improve the accuracy are also important topics of advanced OPC. Traditionally, the OPC correction method consumes a lot of manpower and computing resources. If machine learning is used, the accuracy rate can be greatly improved and the computing time can be reduced.

I have started to use machine learning to develop a hotspot detection method. At present, it can obtain a detection rate of 99.9%. The next stage will focus on using the optimized machine learning model to detect GDS patterns, feed the results back to the OPC verification model, and then modify the OPC recipe so that large-scale corrections can be made efficiently before the circuit is exposed.

研究成果

input
→ **Data Preprocessing**
→ **Classification**
→ **Model Enhancement**
→ **output**
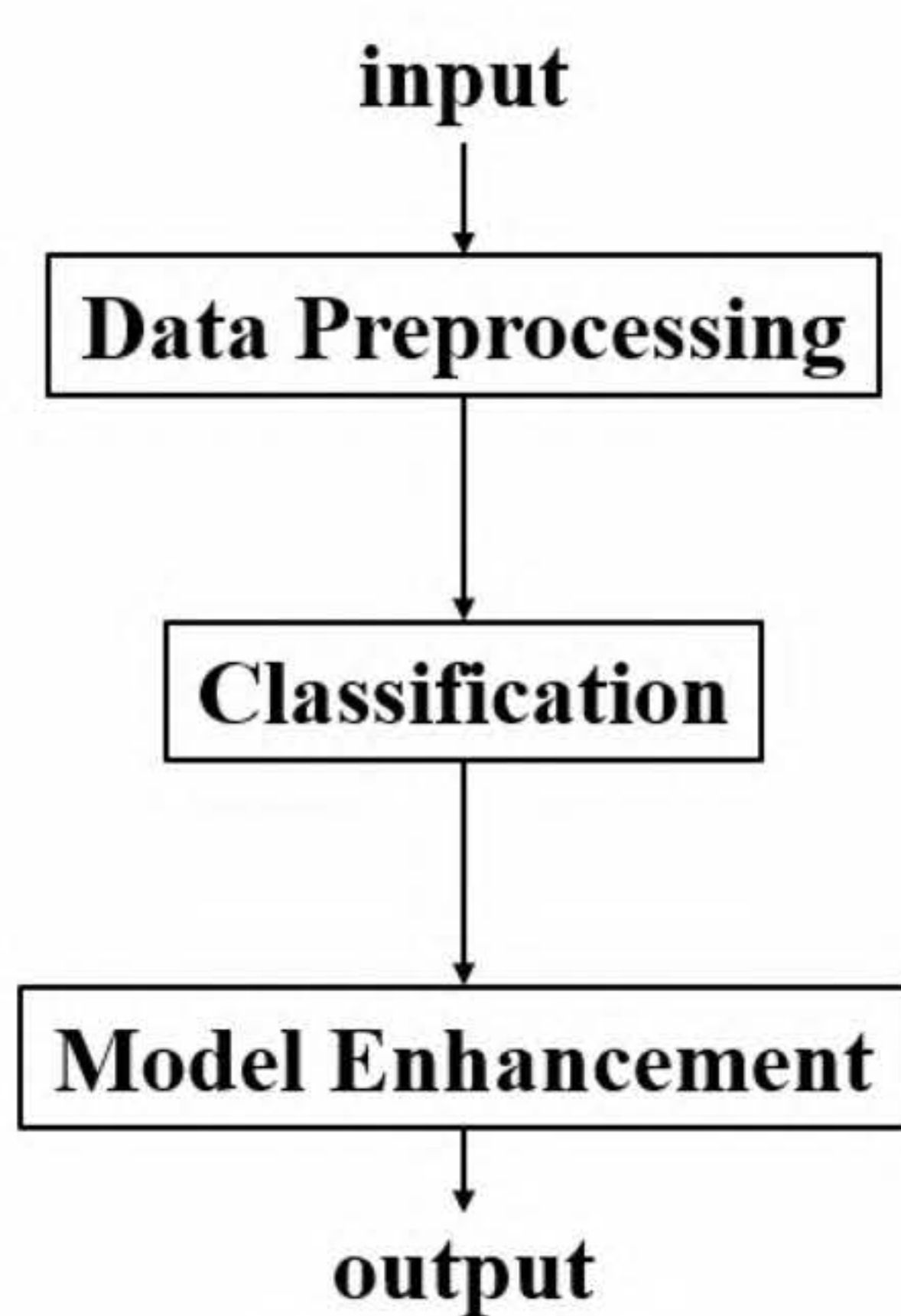
Class imbalance:
- Augmentation (Zoom/contrast)
- Oversampling/Downsampling
- Random-mirror flipping and upsampling
- SMOTE (Synthetic Minority Oversampling Technique)

- Supervised model (ResNet50V2/Xception.etc)
- Semi-supervised model

- Incremental learning

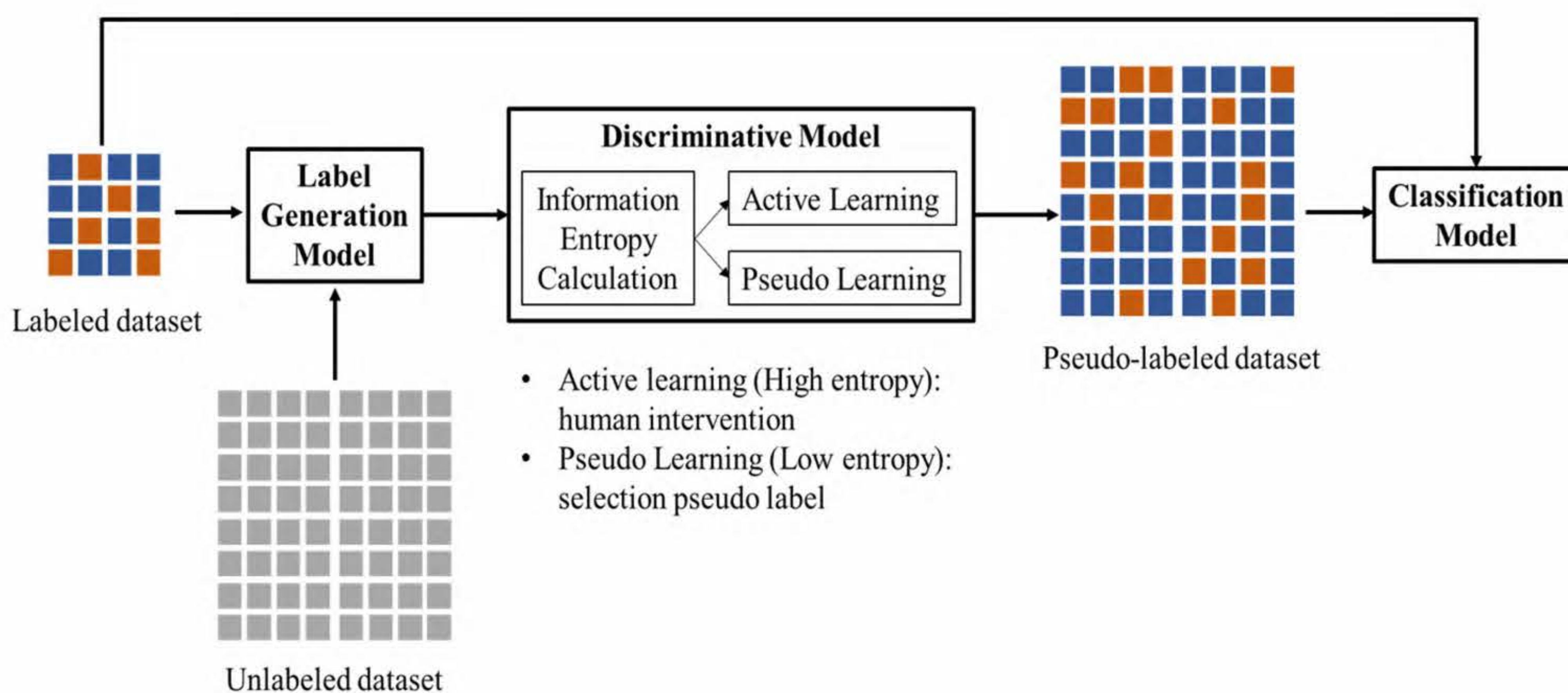Current methods to solve data imbalance include:
- Oversampling/undersampling
- Set a larger loss weight for minority categories
  - Increase the detection rate of minority class
  - Not good for majority class
  - Cause false alarm to increase

A novel two-stage classifier will be proposed to reduce false alarms.

Combined with the NGR images detection results, the ML model is used to directly detect the OPC contour pattern.

Labeled dataset → **Label Generation Model** → **Discriminative Model** (Information Entropy Calculation → Active Learning / Pseudo Learning) → Pseudo-labeled dataset → **Classification Model**

Unlabeled dataset

- Active learning (High entropy): human intervention
- Pseudo Learning (Low entropy): selection pseudo label

Semi-supervised learning architecture flowchart

研究生活及心得

由衷感謝中技社獎學金給予的肯定，我想將這份榮耀與我的指導教授及實驗室夥伴分享。感謝我的指導教授-李克強老師從我專題生時期給予的協助及教導，培養我撰寫學術文章的技巧，並帶領我從事半導體科技公司龍頭之一的產學合作計畫，使我在機器學習方面的研究能與實務接軌。此外，感謝國科會計畫的支持，讓我們能在研究領域發揮所長。期許自己未來能學到更多的經驗與知識，在研究領域有新的突破，對社會貢獻所學、發揮所長。