Assisting human experts in the interpretation of their visual process: A case study on assessing copper sheet surface adhesive potency

Tristan Hascoet
Kobe University
tristan@people.kobe-u.ac.jp

To-san
Kobe University
xxx

Sachiko Nakamura
xxx
XXX
xxx

Tomoko Hayashi
xxx
XXX
xxx

Mari Sugiyama
XXX
XXX
xxx

Yasuo Ariki
XXX
XXX
xxx

Tetusya Takiguchi
XXX
XXX
xxx

Abstract

Deep Neural Networks are often though to lack interpretability due to the distributed nature of their internal representations. In contrast, humans can generally justify, in natural language, for their answer to a visual question with simple common sense reasoning. However, human introspection abilities have their own limits as one often struggles to justify for the recognition process behind our lowest level feature recognition ability: for instance, it is difficult to precisely explain why a given texture seems more characteristic of the surface of a finger nail rather than plastic bottle. In this paper, we showcase an application in which deep learning models can actually help human experts justify for their own low-level visual recognition process: We study the problem of copper sheet adhesive potency assessment in which an expert material scientist is tasked with assessing the adhesive potency of a copper sheet from microscopic pictures of its surface. Although highly trained material experts are able to qualitatively assess adhesive potency, they are often unable to precisely justify for their decision process. We present a model that, under careful design considerations, is able to provide visual clues for human experts to understand and justify for their own recognition process. Not only can our model assist human experts in their interpretation of the surface characteristics, we show how this model can be used to test different hypothesis of the copper surface response to different manufacturing processes.

1. Introduction

Humans are experts in communicating the reasoning process behind their answer to visual questions. For instance, on typical Visual Question Answering (VQA) samples, human annotators are often able to precisely justify in natural language the reason behind their answer to a certain visual question using simple common sense reasoning.

spatial and causal relationships

In contrast, deep Learning models are often viewed as black box predictors lacking interpretability in the sense that existing tools often fail to explain the decision making process behind the models predictions. For instance, a deep learning model trained end-to-end on a VQA dataset may be able to provide the same answer as its human counterpart, but .

While it is true that humans can justify for their answers on high level reasoning tasks, humans also often fail to explain the process behind their low-level feature recognition ability: for example, precisely defining the nature of a specific texture (what are the defining features of a plastic or wooden surface?) and a low-level part attributes exhibiting large intra-class variations (what is the defining features of a "leg" or a "wing"?). Humans constantly perform such low-level visual recognition tasks while being unable to precisely justify for their recognition process.

In this paper, we present one very practical instance of such situation in the micro-processor chip industry, in which expert material scientists are tasked with assessing the adhesive potency of copper sheets. We propose a model that, under careful design considerations, is able to provide visual clues for human experts to understand and justify for their own decision process.

Our model is carefully designed so that a subset of its internal representations carry semantically meaningful Information that can be visualized and easily interpreted by the human experts, confirming their intuitions and eventually

1

shading light on their own decision processes and biases.

xxx NEED INDUCTIVE BIAS MENTION xxx

We then show how these semantically meaningful representations can be used To formulate and validate hypothesis on the physical phenomenon underlying the observed signal; i.e. the degradation of the copper sheet surface as a function of time and atmospheric conditions.

In essence, the argument this paper is aiming for is as follows: although deep learning models lack the "common sense reasoning abilities and the powerful formalism of natural language to communicate and justify for their decision making process, they can provide powerful tools to shed some light into the low-level recognition process they share with humans [?], and for which human introspection often fails to provide convincing explanations.

In practice, the contribution of this paper is as follows: - We formalize a segmentation procedure for statistical surface based on probabilistic lazy label segmentation framework - We show how the model generalization can be used as a proxy metric to quantify the validity of an hypothesis. - We propose a simple model of material surface degradation through time and decompose a model architecture into a recognition module and an hypothesis module that directly learns to formulate the most pertinent hypothesis based on

The remainder of this paper is organized as follows: In Section 2, we present background information on our task: we detail the practical stakes and problem definition, etc. In Section 3, we detail our approach and etc. Section 4 relates our work to existing research Section 5 presents the experiments setting and discuss our results and Section 6 concludes this paper.

2. Background

A wide variety of electronic products, such as smartphones, tablet computers, PCs, cars and TVs, help to make our life convenient and comportable. Almost all of these electronic devices that are computerized are driven by microprocessor chips, into which large numbers of tiny transistors are integrated. Future development of the electronics technology will require chips to be even smaller, thinner and more advanced.

As the foundation for the construction of micro-chips, electronic substrate is made of copper wires through which electricity flows from copper and insulators (resin, resist, prepreg, etc.) that separates the wires, and has a multi-layered structure in which several layers overlap.

[FIG]

The bulit-in copper microstructures could be easily peeled off from resin due to weak adhesion, sometimes just because of the impact of dropping the smartphone or the heat generated when running heavy calculations for machine learning, and this will cause the smartphone broken or electrical short circuits.

Therefore, it is necessary to increase metal-to-resin bonding in order to enhance the reliability of these electronic devices.

There are three conventional methods to strengthen metal-to-resin bonding, which are the use of adhesive (to fasten metal and resin togather), physical adhesion (such as buffing and scrubbing) and chemical adhesion. Here we should notice some facts that adhesion strength depends on surface roughness, high roughness provides relatively high adhesion strength, while low roughness provides relatively low adhesion strength. In our research, we concentrated on chemical adhesion enhancing technology, which adheres metal and resin by chemically treating the metal surface.

The surface treatment technology offers added value by dissolving metal. With this techonology, chemicals are used to produce unique, super-fine surface roughness by soft-etching copper surfaces, enhancing copper to resin adhesion drastically. The chemicals creat the hidden "surfaces" that form the boundaries between metal and resin, dissolve and modify metal surfaces to provide for an enhanced adhesion between the various layers which compose a printed wire board. As the electronic substrate develops smaller and thinner, chemicals are used for a wide range of electronic substrates especially those requiring high reliablity.

In manufacturing process, a large copper clad laminate is transported on a conveyor transfer line, and chemical solutions are sprayed onto the flowing copper clad laminate for pretreatment. However, the degradation process will occur on the surface of these pretreated copper substrates under certain conditions (such as atmosphere or temperature), which means the quality of the surface is getting worse so that metal-to-resin bonding is getting weaker.

To ensure the protection of the device, it is necessary to monitor evolution of the copper surface considering high temperature. In our research, the process of degradation was captured by the microscopes every week, taken in different degrees and scales.

We demonstrated the performance of our experiment on the segmentation of the pixel-wise classification of Scanning Electron Microscopy(SEM) images of copper surface. For now, we have SEM images of 14 weeks, with 50 images taken in different places for each week. In order to augment our dataset, the microscope was set in degrees of 0 (vertical) and 45, and scales of 1000 and 3000, separately.

So that, there are 4 sets of datasets due to different angles and magnifications, and each dataset consists of 50×14 images with size of 960×1280.

For each dataset, we divided it into three parts: training data, testing data, and validation data. For each week, we took out 40 images for training data, 5 for testing data and 5 for validation data. Thus, we had 40×14 images for training data, 5×14 images for testing data, and 5×14 images for validation data.

The condition of the experiment was controlled at a certain temperature. And the pretreated copper clad laminate was placed in the experimental environment without any other operations, only to be taken out when taking photos.

At present, it tends to use expert viewing methods to assess the quality of surface pictures. Material experts can judge whether the pretreated surface is of adequate quality by looking at these pictures directly. Although the material expert can roughly assess the quality of the copper surface, he can't precisely justify his decision, arguing that his decision is made from a subjective synthesis from his experience, looking at plenty of pictures of different copper surface.

3. Method

3.1. Framework

xxx

3.2. Recognition module

xxx

3.3. Hypothesis formulation

xxx

3.4. Hypothesis module

xxx

4. Related Work

Although deep neural networks have achieved a great success on a variety of challenging visualization tasks in recent years, our understanding of how these neural network models is far from enough to interpret. The pursuit of figuring out what is learned for each layer of models and how trained neural network models really "think" never stops.

Jason Yosinski et al. provide two useful tools for visualizing and interpreting neural nets. One is to visualize the activations produced on each layer of a trained convnet as it processes an image or video, and the other enables visualizing features at each layer of a DNN via regularized optimization in image space. Shan Carter et al. create an explorable activation atlas of features the network has learned, by using feature inversion to visualize millions of activations from an image classification network, which can reveal how the network typically represents some concepts. There is a new dataset called Visual Question Answering(VQA) containing open-ended questions about images.These questions require an understanding of vision, language and commonsense knowledge to answer. Trevor Darrel et al. build models (such as Pointing and Justification-based explanation model) to explain their decisions, generating convincing explanations. TensorFlow provides an attention-based model, which enables us to see what parts of the image the model focuses on as it generates a caption.

Deep learning has showed remarkable success on many important learning problems in chemistry, drug discovery, biology and materials science.

A general framework for supervised learning on graphs called Message Passing Neural Networks (MPNNs) is widely used on neural computation for predicting quantum states of molecules. Drug discovery is also benifiting from the development of artificial intelligence technology, which is automating the invention of new chemical entities and the mining of large databases, in drug design and molecular medicine field. Deep learning approaches such as Deep-Mind's AlphaFold are helpful for scientists to deal with the "protein folding problem", not only predicting the intricate 3D structure of a protein but also predicting the physical properties of a protein structure, and are making significant progress on one of the core challenges in biology. In the field of materials science, deep neural networks have also been receiving inreasing attention and have achieved great improvements, for example, in material property prediction and new materials discovery for batteries.

An other point we should pay attention to is different labelling strategies in conventional supervised learning. Typically speaking, it's often assumed that each instance is associated with one single label. However, there should usually be more than one labels for one instance in real-world tasks, if it is multi-label learning.

Eugene et al. propose a semi-supervised framework that employs image-to-image translation between weak labels. Yuxing Tang et al. build a similarity-based knowledge transfer mode trying to investigate whether knowledge about visual and semantic similarities of object categories can help improve the peformance of detectors trained in a weakly supervised setting.

The most related work with ours is Rihuan Ke's[*] in which they present a semi-supervised learning strategy for segmentation with lazy labels and develop a multi-task learning framework to integrate the instance detection, separation and segmentation within a deep neural network.

Our work aim to merge segmentation with lazy labels into scientific discovery and provide material experts with strong supports for interpretable decision.

5. Experiments

xxx

5.1. Recognition results

xxx

5.2. Hypothesis testing

xxx

5.3. Hypothesis learning

xxx

6. Conclusion

xxx

References