

## ディスプレイ画質検査における AI 応用

○尾河 雄太<sup>†</sup>, 王 佳璽<sup>†</sup>, 黒木 健司<sup>†</sup>, 三谷 洋之<sup>†</sup>, 劉 偉<sup>†</sup>, 浅野 敏郎<sup>‡</sup>

†: 株式会社ファースト 検査システム開発本部 検査システム技術グループ

‡: 広島工業大学

Y.Ogawa@fast-corp.co.jp

概要：近年，LCD・OLED 薄型テレビやスマートフォンの普及に伴って高精細パネルの生産量が増加している．生産量の増大による検査員不足や検査要員の個人差に依存しない安定な検査を実現するため，検査の自動化への要求が高まっている．実際のディスプレイ検査装置では欠陥を検出するだけでなく欠陥分類まで細かく要求されることが多い．この欠陥分類精度はディスプレイ生産の歩留まりに直接影響を与える．そこで本報告では，従来手法のみでは困難な欠陥分類及び過検出抑制を目指し Convolutional Neural Network を用いた自動分類の可能性を検証した．

<キーワード>CNN, ディスプレイ, 画質検査, ムラ

### 1. はじめに

近年，外観検査等の検査における欠陥の検出・分類に対する AI の実応用が盛んに行われている<sup>[1]</sup>．実際に市場からも AI を利用した装置を求められることが多くなっている．この理由の一つとして，複雑な画像処理設定を行うことなく検査性能を調整できるという点が挙げられる．これは装置メーカーにとってもパラメータ調整の負担を軽減できるため，双方に利点がある．

実際のディスプレイ画質検査では，欠陥検出を行うために様々なアルゴリズムが研究されている<sup>[2]</sup>．しかし，欠陥検出を行うことが可能であったとしても分類することが難しい欠陥も多く存在する．欠陥分類はパネルメーカーによっては数十種類存在する場合もあり，複数の特徴量を組み合わせるだけでは全てを分類することは実質不可能である．また，代表的な欠陥である黒点欠陥いわゆるドット抜けは個数や連結数などにより不良レベルを切り分ける必要があり，分類精度はパネル生産の歩留まりに大きく影響する．さらには，線欠陥を点欠陥として検出してしまうケースやパネルのアクティブエリア端部分を欠陥として誤認識してしまうケースも存在する．ディスプレイ画質検査装置では未だにこのような課題が多く存在する．

本報告では，当社のフラットパネルディスプレイ検査装置と Convolutional Neural Network (以下 CNN) を組み合わせて通常の特徴量では分類が容易ではない欠陥の AI 自動分類，及び過検出抑制を目的に実応用の可能性について検討した．

### 2. フラットパネルディスプレイ検査装置

図 1 にフラットパネルディスプレイ検査装置の構成図を示す．フラットパネルディスプレイ検査装置はカメラ，光学系，画像処理装置，ホスト PC から構成されている．カメラと 1 対 1 に対応した画像処理装置にて欠陥検出を行い，欠陥情報はホスト PC に集約され，ここで欠陥の分類と検査結果の判定を行う装置構成となっている．本報告に使用したカメラは Vieworks 製の VP-71MC-M4E0 (7100 万画素，モノクロカメラ)，レンズは Pentax SMCP-FA645(55-110mm, F5.6) を使用した．

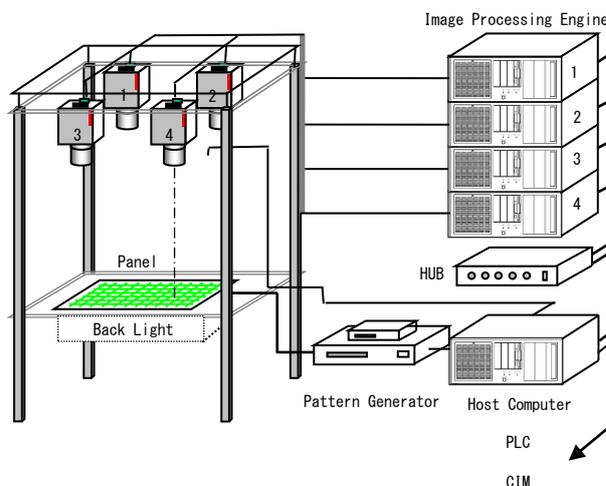


図 1 ディスプレイ検査装置構成例

### 3. 画質検査の欠陥種別

ディスプレイの画質検査では様々な欠陥の検出及び欠陥分類が必要である。大判ガラスを切断し最初の簡易点灯を行う工程、いわゆる Cell 工程における主な欠陥の種類を図 2 に示す。図 2 の様な大分類を見る限りでは欠陥の種類はあまり多くない様に見えるが、実際には各欠陥に対してさらに数十もの欠陥の種類が存在する。

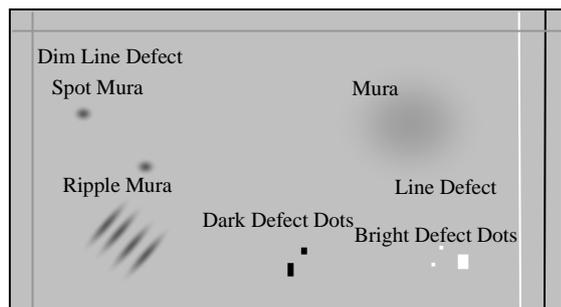


図 2 Cell 工程の画質検査の主な欠陥種類

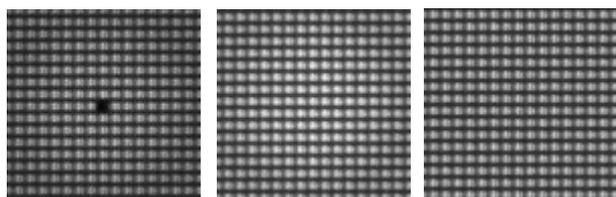
### 4. 学習モデルの検討

従来のフラットパネルディスプレイ検査装置のアルゴリズムだけでは分類が困難な欠陥を対象に検証を行う。フラットパネルディスプレイ検査装置で検出された欠陥座標を中心に画像を 96×96 Pixel のサイズで切り出した画像を使用した。

#### 4.1. ムラ核有り・核無し分類

ムラ欠陥において点欠陥とムラ欠陥の特徴も持ち合わせた欠陥を対象とした。対象の欠陥を図 3 に示す。この欠陥はムラ欠陥の中心に黒点が存在することが特徴的で、周囲に発生するムラ欠陥の色と濃さは不定である。この欠陥を学習させる際には「核有り」、「核無し」、「欠陥無し」の 3 クラスとした。

学習フェーズでは、学習データとして 448 枚、評価データを 449 枚としランダムにサンプリングした。学習は 96×96 Pixel のサイズで Batch Size を 64、エポック数を 100 回に設定して学習を行った。



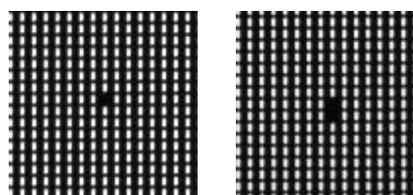
(a) 核有ムラ (b) 核無しムラ (c) 欠陥無し  
図 3 ムラ欠陥画像例 (疑似欠陥)

#### 4.2. 黒点欠陥分類

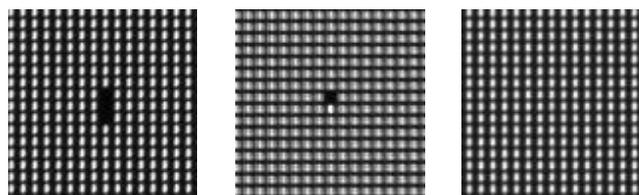
黒点欠陥はディスプレイの画質検査において、代表的な欠陥であり、分類においては最も煩雑な欠陥でもある。この理由として、欠陥の個数及び連結数によって不良レベルを切り分ける必要があるためである。

AI を用いてこれらを容易に分類することができるのかを検証した。対象の欠陥を図 4 に示す。この欠陥を学習させる際には「単黒点」、「2 連結」、「3 連結以上」、「明暗連結」、「欠陥無し」の 5 クラスとした。

学習フェーズでは学習データとして 1243 枚、評価データを 1244 枚としランダムにサンプリングした。学習は 96×96 Pixel のサイズで Batch Size を 64、エポック数を 100 回に設定して学習を行った。



(a) 単黒点 (b) 2 連結



(c) 3 連結以上 (d) 明暗連結 (e) 欠陥無し

図 4 黒点欠陥画像例 (疑似欠陥)

#### 4.3. 学習モデル

筆者らが作成した学習モデルの構造を表 1 に示す。Model\_1 は独自に作成したモデルとなっており、コンボリューション層の次にバッチ正規化を入れているモデルである。このモデルを作成した理由として、小規模なモデルでどの程度の正解率を出すことができるのかを評価するためである。

Model\_2 は 1 枚の画像に対して、2 通りに枝分かれさせ最終的に結合し双方の MAX 値により最終判定を行う構造となっている。具体的には、最初の Convolution の際に 3×3 のカーネルサイズの場合と 7×7 のカーネルサイズで分岐させている。これは、核有りムラが狭域的な特徴と広域的な特徴を持ち合わせているため、カーネルサイズを大小の 2 つにすることで認識率が上昇するのではないかと推測し作成した。

学習に使用した PC のスペックは CPU が Intel(R)Xeon(R)E5-1650v4、RAM が 32GB、GPU は

Quadro RTX4000(VRAM 8GB)である. なお, 学習モデルの作成は Sony 社製の Neural Network Console を使用した.

表 1 学習モデルの構造

(a) Model\_1 (b) Model\_2

Input96	Input96	
Conv3-64	Conv3-64	Conv7-64
BatchNormalization	BatchNormalization	BatchNormalization
ReLU	ReLU	ReLU
Max Pooling	Max Pooling	Max Pooling
Conv3-128	Conv3-128	Conv3-128
BatchNormalization	BatchNormalization	BatchNormalization
ReLU	ReLU	ReLU
Max Pooling	Max Pooling	Max Pooling
Conv3-256	Conv3-256	Conv3-256
BatchNormalization	BatchNormalization	BatchNormalization
ReLU	ReLU	ReLU
Max Pooling	Max Pooling	Max Pooling
Conv3-512	Conv3-512	Conv3-512
BatchNormalization	BatchNormalization	BatchNormalization
ReLU	ReLU	ReLU
Max Pooling	Max Pooling	Max Pooling
Conv3-512	Conv3-512	Conv3-512
BatchNormalization	BatchNormalization	BatchNormalization
ReLU	ReLU	ReLU
Max Pooling	Max Pooling	Max Pooling
Conv3-256	Conv3-256	Conv3-256
BatchNormalization	BatchNormalization	BatchNormalization
ReLU	ReLU	ReLU
Affine	Maximum	
SoftMax	Affine	
	SoftMax	

## 5. 学習モデルの検討結果

表 2 に結果を示す. 参考として Vgg<sup>[3]</sup>及び ZFNet<sup>[4]</sup>の結果も記載する. ZFNet 以外ではほぼ近い正答率となったが, 核有りムラ用に作成した Model\_2 の結果がムラ分類においては一番良い結果となった. また黒点分類においては Vgg16 が一番良い結果を示した. 比較的小規模な Model\_1 と Model\_2 においても Vgg16 と遜色無い精度であることから著者らが作成した学習モデルの有用性を示すことができたと考えられる. また, 規模が小さい学習モデルであるため, 学習時間や推論をする際のタクトタイムに良好な影響を与えるのではないかと考えられる.

以上の結果から, 学習モデルは判定精度の高い Model\_2 を用いてインラインでの検証を行うこととした.

表 2 各モデル比較結果

学習モデル	平均正解率(10回平均)	
	ムラ分類	黒点分類
Model_1	92.4%	94.6%
Model_2	93.3%	94.7%
Vgg11	92.1%	93.9%
Vgg16	91.8%	95.6%
ZFNet	88.6%	93.7%

## 6. インラインテスト

実際の後工程ラインに AI を導入し, 欠陥分類精度及び過検出の抑制効果の検証を行った.

図 5 に検査フローを示す. 画像処理を用いて欠陥を検出するところまでは従来と同様の検査フローとなるが, 欠陥を分類する際に AI 判定を用いる. すべての欠陥に対して推論を行うわけではなく, 分類が困難な欠陥のみ推論を行うフローとなっている. また, AI の分類が困難な欠陥の場合は検査員が再判定を行いデータの更新を行う. AI が分類困難であるという基準は「SoftMax の値が 99%を下回っている時」と定義した.

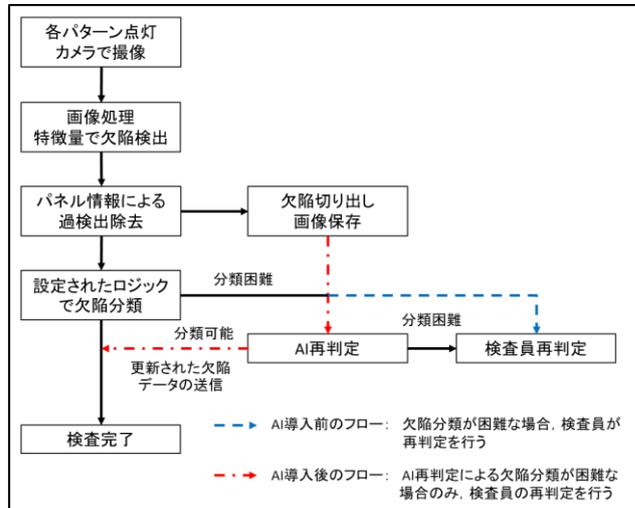


図 5 検査フロー

### 6.1. インラインテスト 1

インラインテスト1ではムラ欠陥を対象に検証を行った. 学習モデルの検討では, 「核有り・無し」を分類したが, 現場から「不良レベルの切り分けを優先したい」との要望があり, 最初に不良レベルの切り分けを検証することにした. なお, 不良レベルとは欠陥の程度を表すものであり, この程度により良品扱いにするか不良品扱いにするかを定めることになる.

インラインテスト 1 の対象欠陥の一例を図 6 に示す。

学習フェーズでは学習データとして 18349 枚を使用した。学習は  $192 \times 192$  pixel を  $96 \times 96$  Pixel にリサイズした上で Batch Size を 64, エポック数を 100 回に設定して学習を行った。クラスは 6 クラスとした。表 3 にクラス分類一覧を示す。表 3 中のアクティブエリア端誤検出は、画像処理で欠陥検出を行う際に検査エリアにズレが生じた場合に発生する。検査エリアにズレが発生する原因としては、なんらかの影響で点灯画面が通常よりも暗い又は明るい状態になった際に検査エリアを検出するためのしきい値を下回ってしまった場合等が考えられる。

表 3 ムラ欠陥クラス分類

クラス
核有りムラ(不良レベル)
映り込み・チャージ等(良品)
核有りムラ(良品レベル)
欠陥無し
アクティブエリア端誤検出
線欠陥誤検出

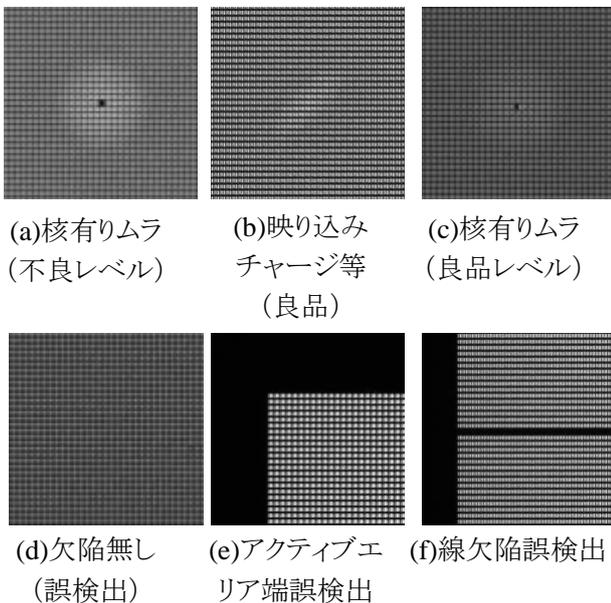


図 6 インラインテスト 1 の対象欠陥の画像(疑似欠陥)

### 6.1.1 インラインテスト 1 結果

結果を表 4 に示す。対象となる欠陥は 2200 個であったが、SoftMax の値が 99% を下回っているものに関しては AI の再判定から除外しているため、最終的に

AI で再判定されているものは 1986 個となった。結果としては、全体の正解率が 99.6% という結果となった。これは筆者らが想定していた数値よりもはるかに高い値を示した結果となった。

表 4 インラインテスト 1 判定結果

クラス	欠陥数	正解数	不正解数	正解率
核有りムラ(不良レベル)	153	153	0	100.0%
映り込み(良品)	44	40	4	90.9%
核有りムラ(良品レベル)	898	898	0	100.0%
欠陥無し	788	786	2	99.7%
アクティブエリア端誤検出	97	96	1	99.0%
線欠陥誤検出	6	5	1	83.3%
	1986	1978	8	99.6%

### 6.1.2 インラインテスト 1 考察

インラインテスト 1 の結果から、従来の特徴量だけでは分類が困難だったものが容易に分類することができていることが読み取れる。特に顕著なものは、映り込みと欠陥無し(誤検出)の場合である。「映り込み」はパネルが鏡となって機構部が映ってしまう現象であるが、楕円状のムラとして見えることが多く、ムラ欠陥との切り分けが非常に難しい。正解率は 90% 程度ではあるが、この程度の正解率であっても過検出の抑制という意味では非常に有効であると考えられる。誤判定をした欠陥画像を確認すると、核有りムラ且つ核がムラの端に存在している様な欠陥を映り込みと誤判定してしまっていることが判明した。通常、核有りムラは図 6(a) の様に核を中心に同心円状に広がるムラのことを指す。そのため、誤判定した欠陥は核有りムラとして学習させていないため、映り込みとして誤判定してしまった可能性がある。これに関しては、再度核有りムラもしくは別の欠陥として学習させることで改善することができると考えている。

また、欠陥無し(誤検出)では、99.7% の分類精度を出すことができている。誤検出をするケースとしては、非常にコントラストが低い欠陥に対して画像処理で強調処理を行う際に本来の欠陥の箇所以外も検出してしまうケースがある。特にディスプレイ自体にシェーディングの様なものがあると誤検出をしてしまう可能性がある。これを 99.7% の精度で抑制することができるとすれば非常に AI は有効な手段であると言える。

## 6.2. インラインテスト 2

インラインテスト 2 では黒点欠陥を対象に分類の検証を行った。この欠陥を学習させる際には 10 クラスとした。表 5 にクラス分類一覧を示す。主な黒点欠陥の画像は図 4 と同様である。

学習フェーズでは学習データとして 6274 枚を使用した。学習は 96×96 Pixel のサイズで Batch Size を 64、エポック数を 100 回に設定して学習を行った。

表 5 黒点欠陥クラス分類

クラス
単黒点
黒点 2 連結
黒点 3 連結以上
欠陥無し
明暗連結
アクティブエリア端誤検出
群黒点
線欠陥誤検出
薄黒点
異物ムラ

### 6.2.1 インラインテスト 2 結果

結果を表 6 に示す。対象となる欠陥は 2193 個であったが、SoftMax の値が 99%を下回っているものに関しては AI の再判定から除外しているため、最終的に AI で再判定されているものは 1905 個となった。結果としては、全体の正解率が 96.9%という結果となった。

表 6 インラインテスト 2 判定結果

クラス	欠陥数	正解数	不正解数	正解率
単黒点	1115	1114	1	99.9%
黒点2連結	117	117	0	100.0%
黒点3連結以上	11	11	0	100.0%
欠陥無し	141	141	0	100.0%
明暗連結	49	49	0	100.0%
アクティブエリア端誤検出	103	84	19	81.6%
群黒点	69	69	0	100.0%
線欠陥誤検出	206	166	40	80.6%
薄黒点	54	54	0	100.0%
異物ムラ	40	40	0	100.0%
	1905	1845	60	96.9%

### 6.2.2 インラインテスト 2 考察

インラインテスト 2 の結果からも従来の特徴量だけでは分類が困難だったものが容易に分類できていることが読み取れる。欠陥無し(誤検出)に関しては判定正解率が 100%となっており、非常に高い正解率を示した。しかし、アクティブエリア端及び線欠陥では正解率が 80%程度と落ち込んでいる。アクティブエリア端の誤検出時の欠陥画像を図 7 に示す。インラインテスト 1 の場合ではアクティブエリア端誤検出に対する分類精度は 99%となっており、インラインテスト 2 の結果と大きく異なっている。これは切り出した欠陥画像のサイズがインラインテスト 1 では 192×192 Pixel であるのに対し、インラインテスト 2 では 96×96 Pixel にしていることが影響していると考えられる。今後は学習前の画像のサイズも詳細に検討する必要があると考えている。

線欠陥誤検出に関しては、アクティブエリアの端を線欠陥として誤検出した際に AI の再判定においても線欠陥として分類されてしまっているケースが多い。これは学習させる際のデータ数が少ないことが問題であると考えている。

黒点 2 連結及び黒点 3 連結以上に関しては、正解率が 100%と非常に高い正解率を示した。しかし、単黒点の正解率が 99.9%となっており、1 枚だけ判定が不正解となったものがある。その欠陥は単黒点を貫通する様に薄い線欠陥が存在するような欠陥であった。このような場合当然欠陥としての不良レベルは線欠陥の方が高いため線欠陥として扱わないといけない欠陥である。将来的にはこのような混在した場合の対策を考える必要がある。



図 7 アクティブエリア端誤検出時の画像(疑似欠陥)

## 7. まとめ

本報告では、従来手法のみでは困難な欠陥分類及び過検出抑制を目指し CNN 用いた自動分類の可能性を検証した。比較的小規模な学習モデルを作成し、インラインテストを行った。その結果、欠陥の種類によっては 80~100%の精度で欠陥分類を行えることが分かった。核有りムラ欠陥の不良

レベルの切り分けに対する分類精度は 100%, 単黒点の分類精度は 99.9%, 黒点 2 連結及び黒点 3 連結以上の分類精度は 100%と非常に高い精度であることが分かった。

以上のことから, ディスプレイ画質検査における AI 応用は可能であり, さらなる検討により従来手法だけでは困難であった欠陥分類を精度良く容易に行うことができるようになる可能性を示すことができた。

今後の課題としては, SoftMax の値が 99%以下の条件により AI の再判定から除外されてしまった欠陥の分類を可能にすることである。本研究の最終的な目標は, 全ての欠陥の AI 自動判定であり欠陥種別のみならず欠陥コード (不良コード) の分類までを行うことである。

**謝辞** 本研究は株式会社ファースト広州事務所の呉明氏の協力を受けており, 記して感謝する。

#### 参考文献

- [1] 池田 泰之, 半澤 雄希, 栗田 真嗣: "AI 技術による外観検査の自動化", ViEW2018, OS3-H2 (2018).
- [2] Kazuyuki ISHIGURO, Toshio ASANO, Takahiro KONDOH and Wei LIU: Inspection of Mura Image Quality by Using Frequency Analysis for Electronic Displays, IDW'16, pp.945-948 (2016).
- [3] Karen Simonyan, Andrew Zisserman: VERY DEEP CONVOLUTIONAL NETWORKS FOR LARGE-SCALE IMAGE RECOGNITION, ICLR2015, (2014), <https://arxiv.org/pdf/1409.1556.pdf>
- [4] Matthew D. Zeiler, Rob Fergus: Visualizing and Understanding Convolutional Networks ECCV.2014, pp.818-833.(2014)

**尾河 雄太**: 株式会社ファースト 検査システム開発本部 検査システム技術グループに所属。ディスプレイ自動検査装置の開発に従事。

**王 佳璽**: 株式会社ファースト 検査システム開発本部 検査システム技術グループに所属。ディスプレイ自動検査装置の開発に従事。

**黒木 健司**: 株式会社ファースト 検査システム開発本部 検査システム技術グループに所属。ディスプレイ自動検査装置の開発に従事。

**三谷 洋之**: 株式会社ファースト 検査システム開発本部 検査システム技術グループに所属。ディスプレイ自動検査装置の開発に従事。

**劉 偉**: 株式会社ファースト 検査システム開発本部 検査システム技術グループに所属。ディスプレイ自動検査装置の開発に従事。

**浅野 敏郎**: 株式会社日立製作所を経て, 1999 年より広島工業大学勤務。現在, 広島工業大学名誉教授。ディスプレイの画質評価, 検査の研究に興味をもつ。