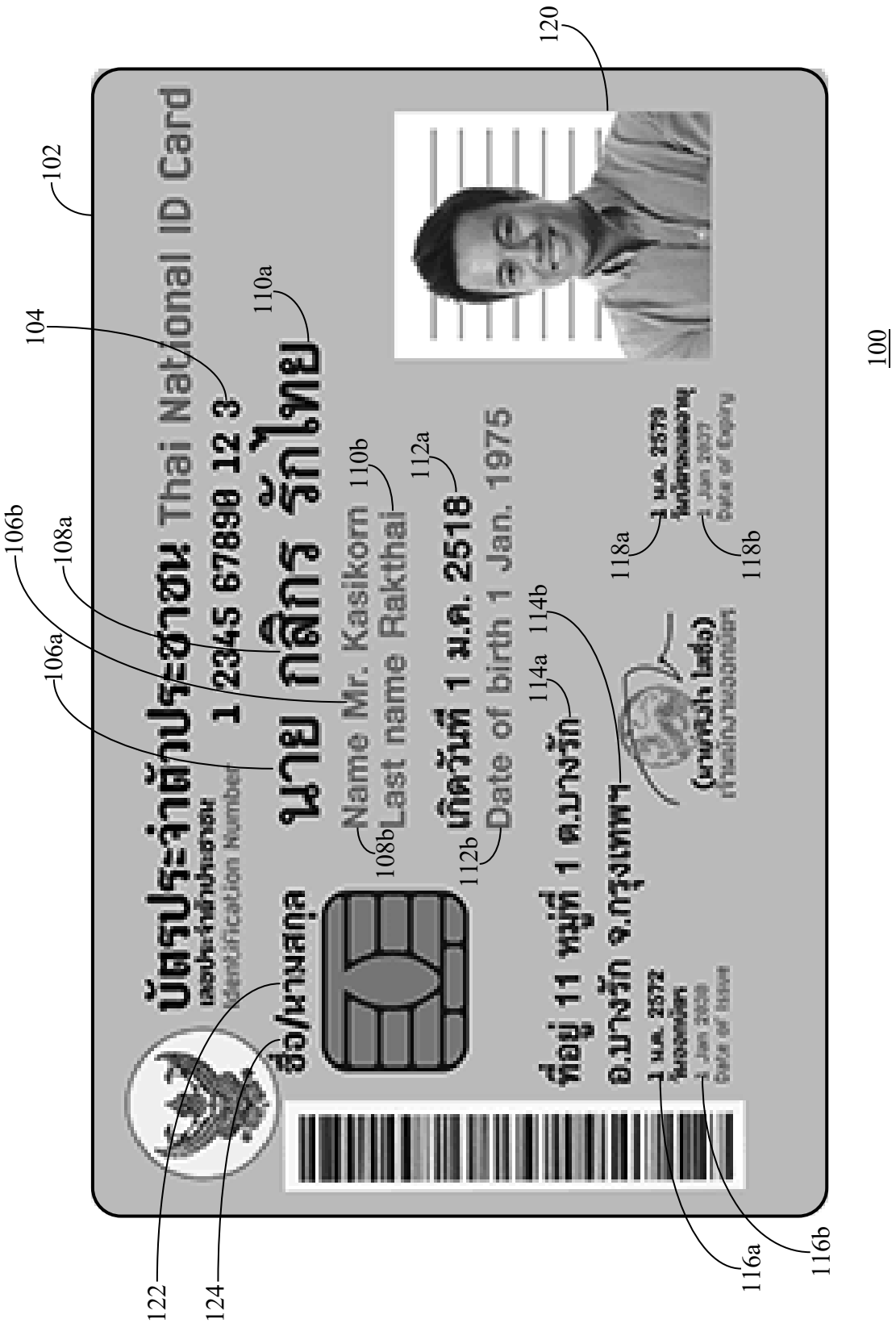
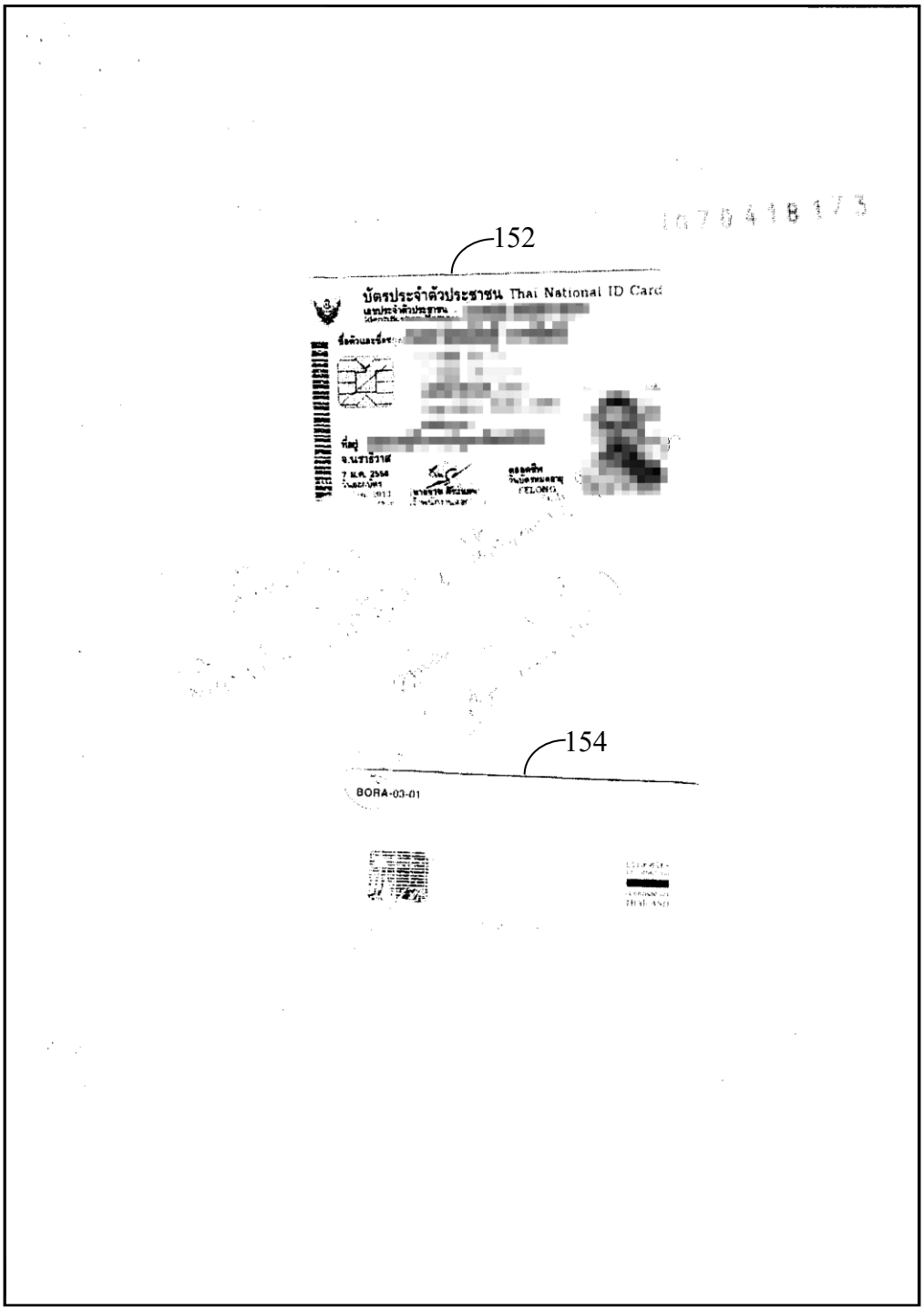

บัตรประจำตัวประชาชน Thai National ID Card
1 2345 67890 12 3
นาย กสิกร รักไทย
Name Mr. Kasikom
Last name Rakthai
เกิดวันที่ 1 ม.ค. 2518
Date of birth 1 Jan. 1975
ที่อยู่ 11 หมู่ที่ 1 ต.บางรัก
อ.บางรัก จ.กรุงเทพฯ
(บางกอก Lado)
102
104
106a
106b
108a
108b
110a
110b
112a
112b
114a
114b
116a
116b
118a
118b
120
122
124
100

FIG. 1A



150

FIG. 1B



FIG. 2A



FIG. 2B

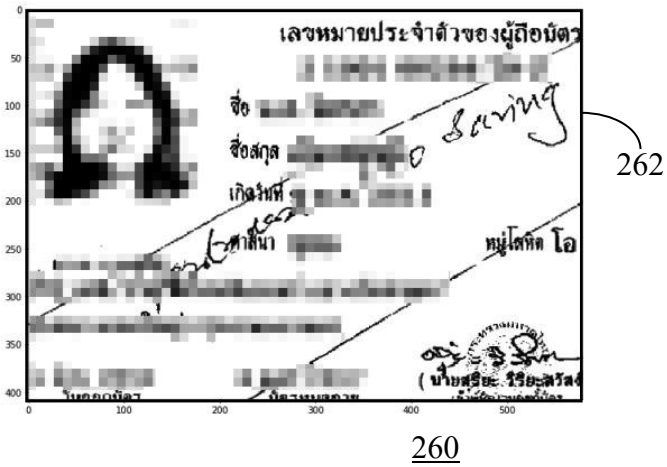


FIG. 2C

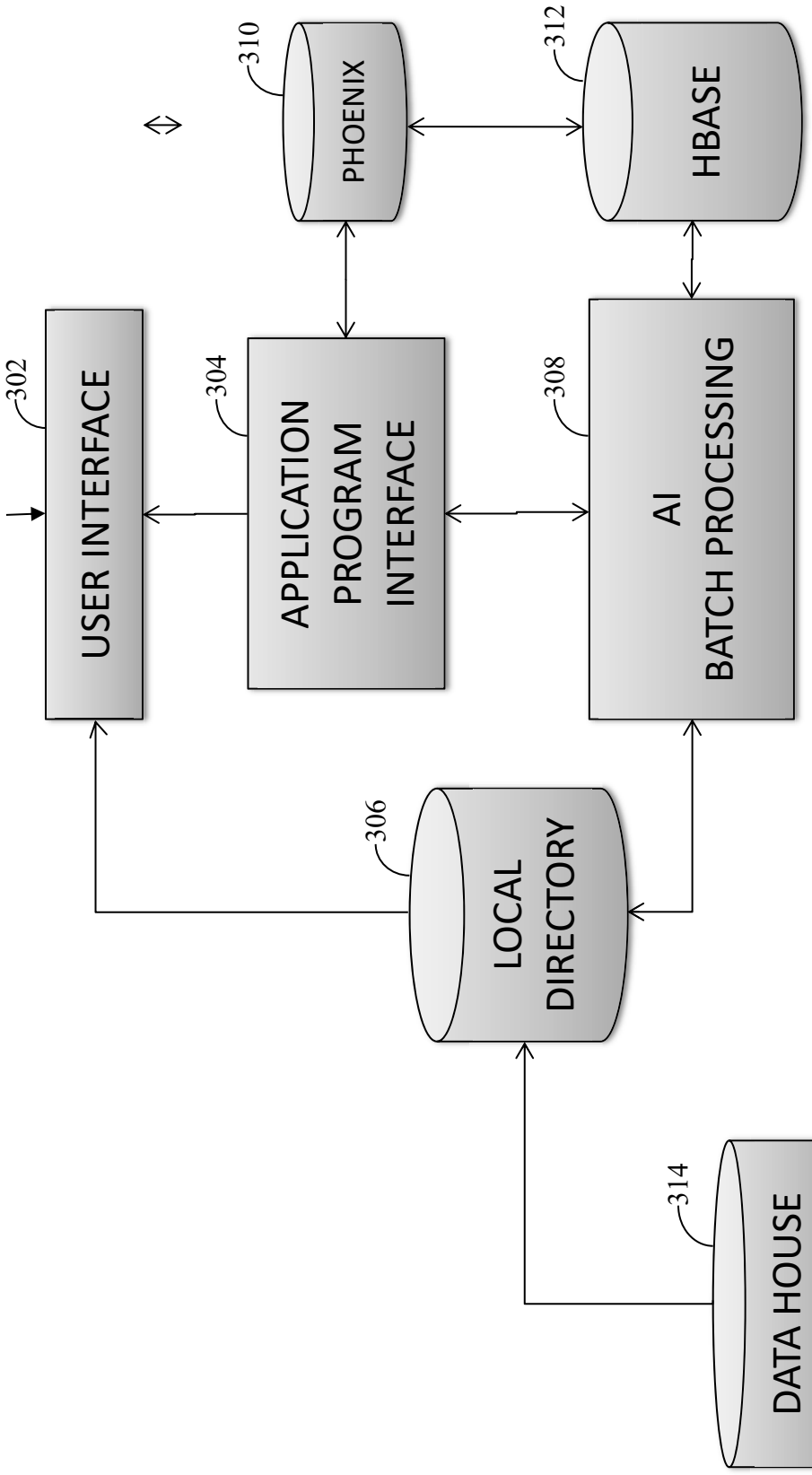

310
PHOENIX
312
HBASE
302
USER INTERFACE
304
APPLICATION PROGRAM INTERFACE
308
AI BATCH PROCESSING
306
LOCAL DIRECTORY
314
DATA HOUSE
300

FIG. 3

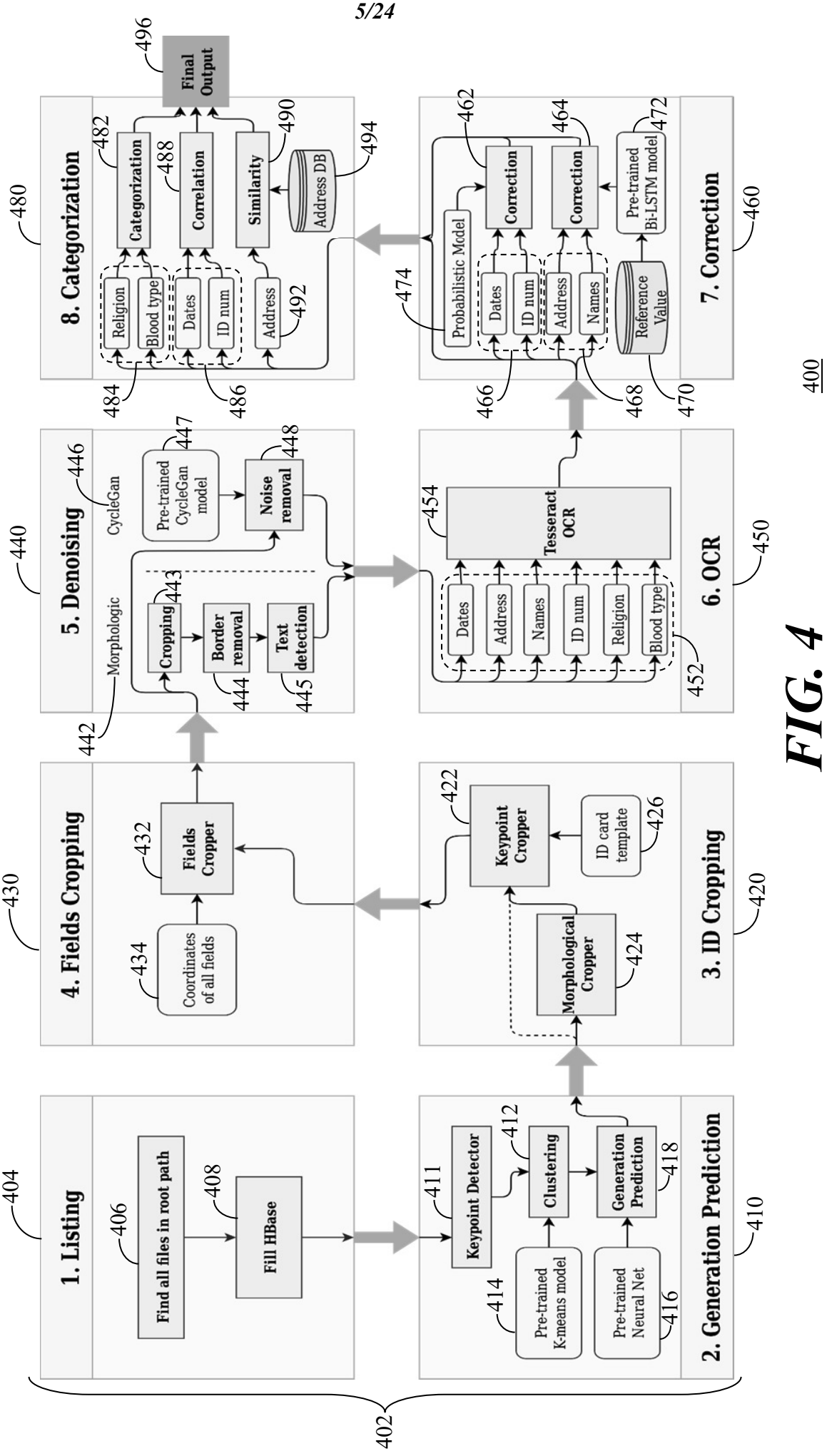

1. Listing
404
Find all files in root path
406
Fill HBase
408
2. Generation Prediction
410
Keypoint Detector
411
Clustering
412
Generation Prediction
418
Pre-trained K-means model
414
Pre-trained Neural Net
416
3. ID Cropping
420
Keypoint Cropper
422
Morphological Cropper
424
ID card template
426
4. Fields Cropping
430
Fields Cropper
432
Coordinates of all fields
434
5. Denoising
440
Morphologic
442
Cropping
443
Border removal
444
Text detection
445
CycleGan
446
Pre-trained CycleGan model
447
Noise removal
448
6. OCR
450
Tesseract OCR
454
Dates
Address
Names
ID num
Religion
Blood type
452
7. Correction
460
Probabilistic Model
474
Correction
462
Correction
464
Pre-trained Bi-LSTM model
472
Reference Value
470
466
468
8. Categorization
480
Religion
Blood type
Dates
ID num
Address
492
Categorization
482
Correlation
488
Similarity
490
Address DB
494
Final Output
496
484
486
402
400

FIG. 4

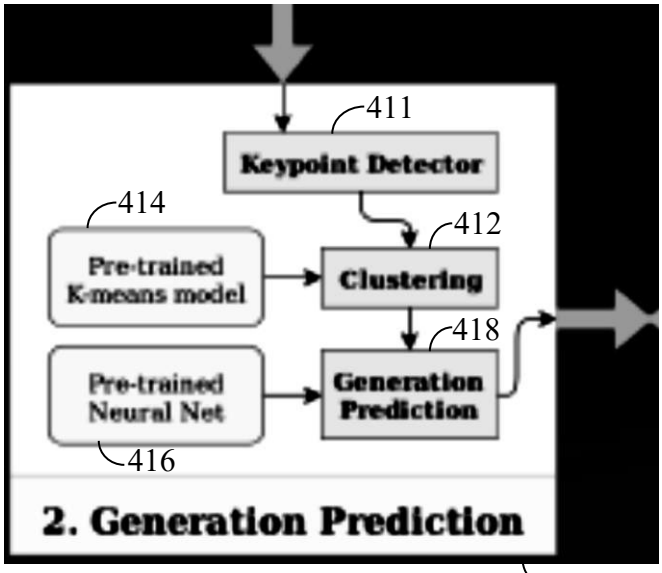

411
Keypoint Detector
414
412
Pre-trained K-means model
Clustering
418
Pre-trained Neural Net
Generation Prediction
416
2. Generation Prediction
500
410

FIG. 5A

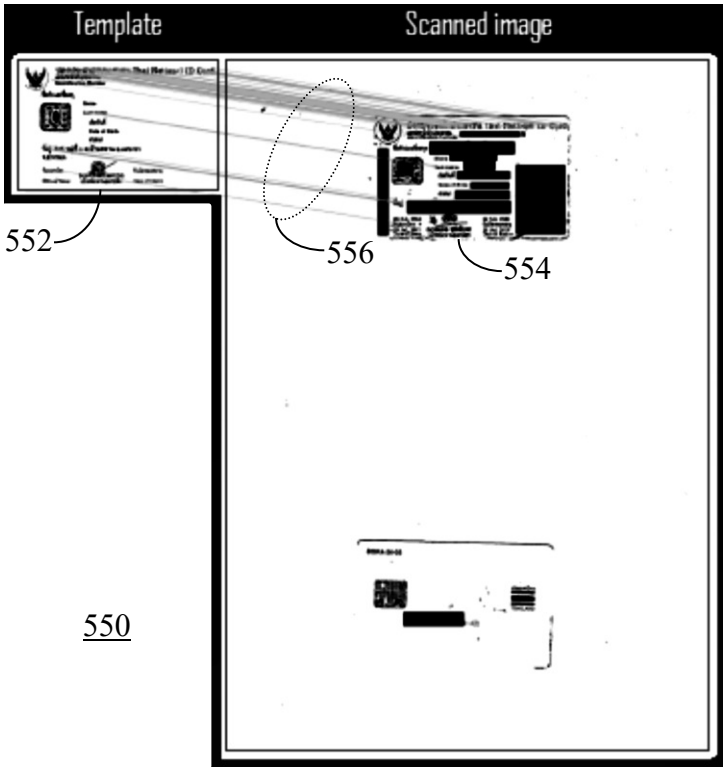
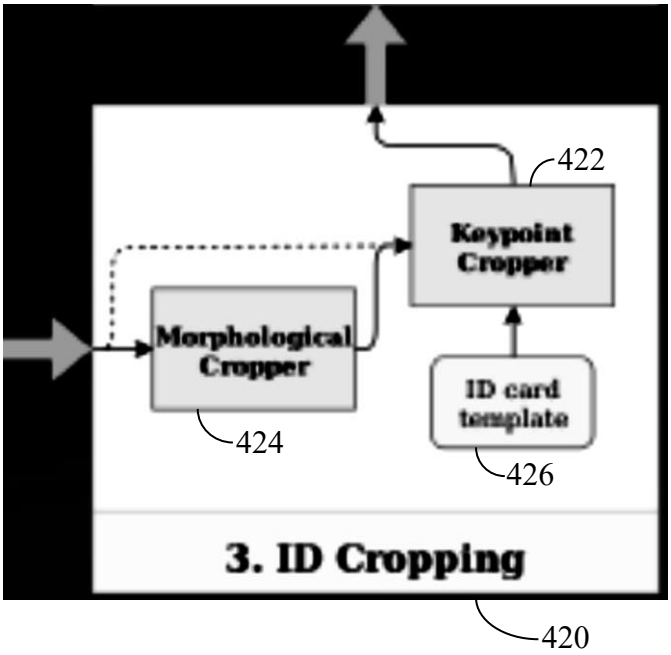

Template
Scanned image
552
556
554
550

FIG. 5B



600

FIG. 6A



610

FIG. 6B

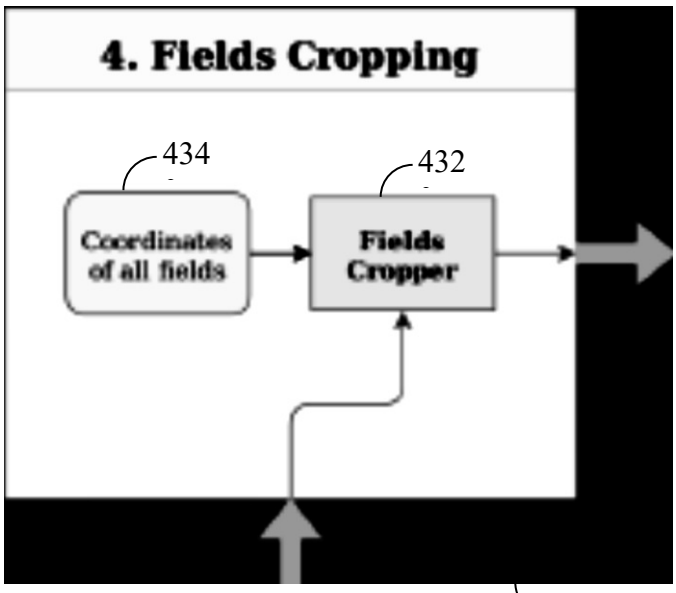

4. Fields Cropping
434
432
Coordinates of all fields
Fields Cropper
700
430

FIG. 7A

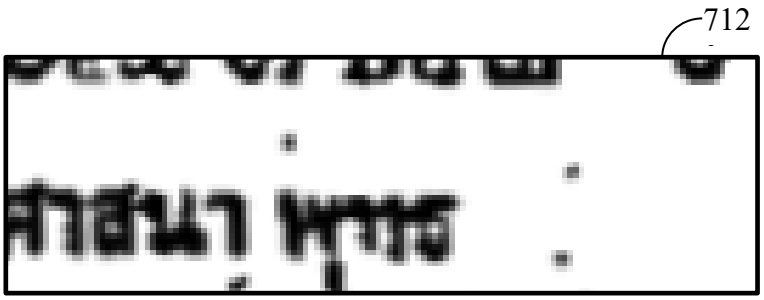
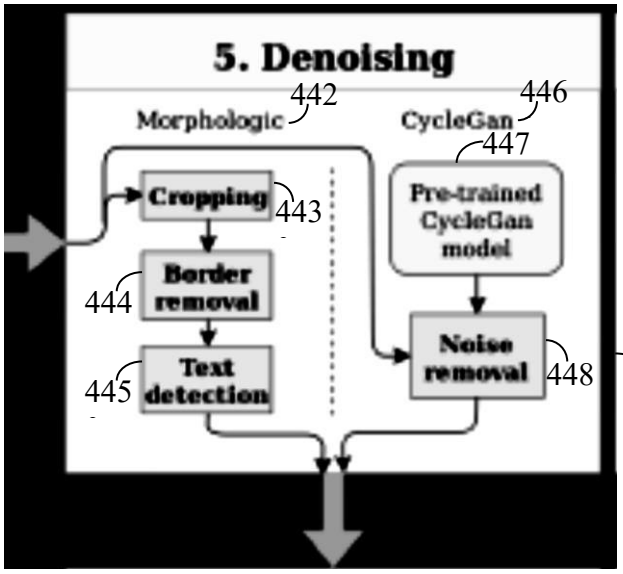

712
710

FIG. 7B



800

FIG. 8A

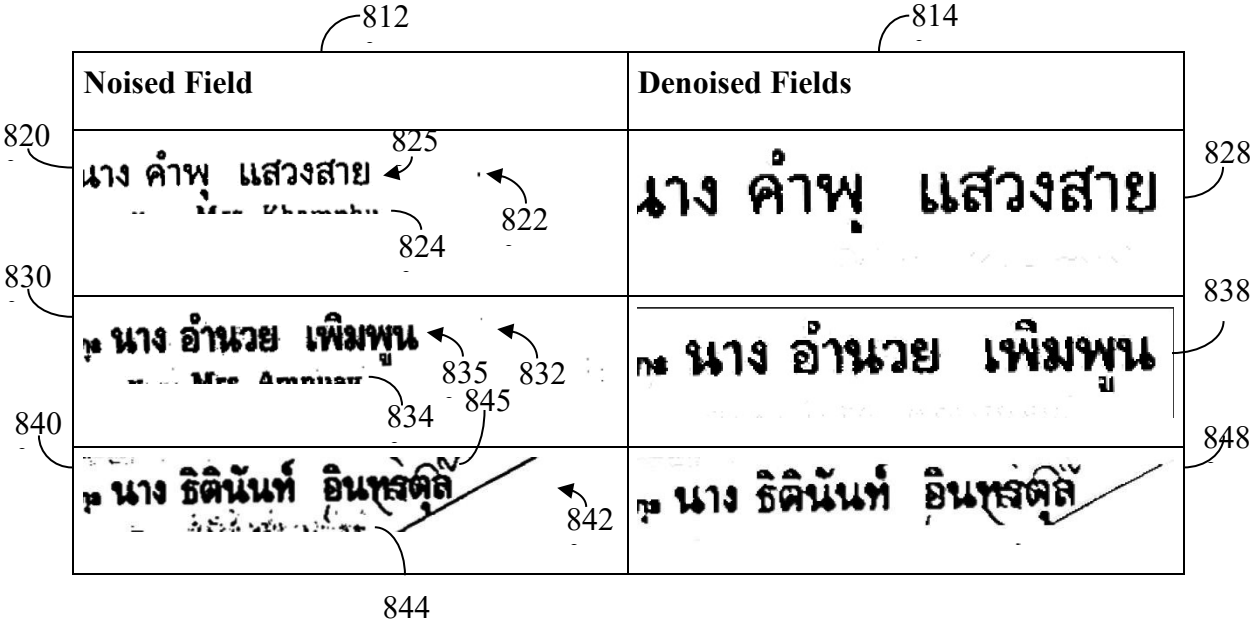


FIG. 8B

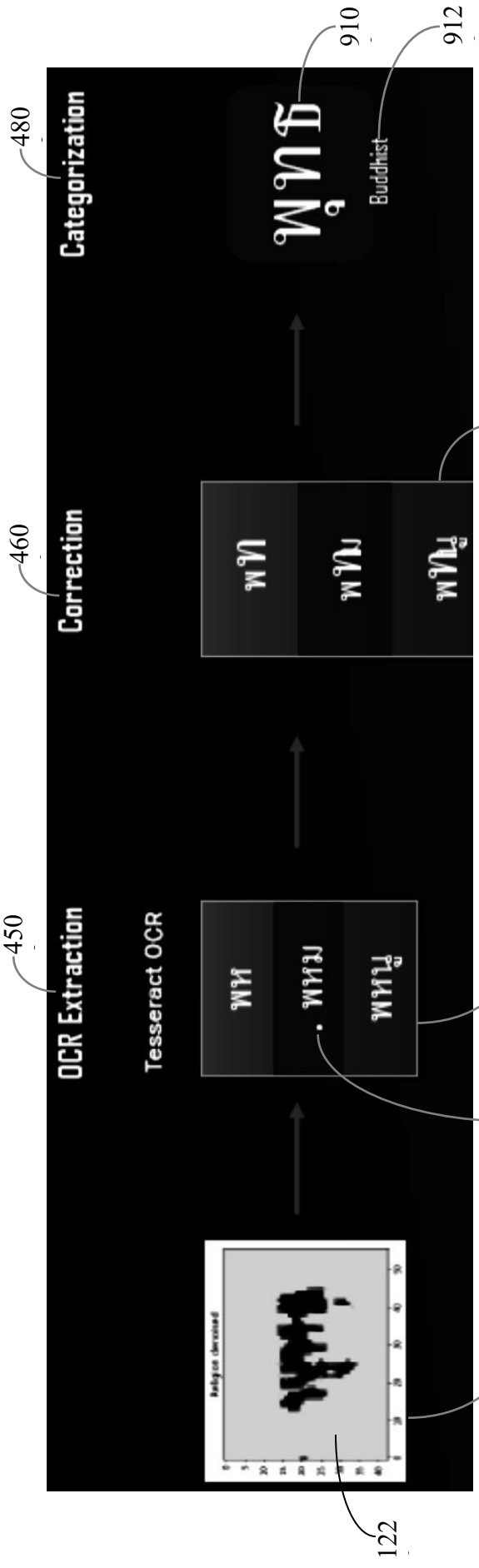

480
Categorization
910
912
Buddhist
460
Correction
908
450
OCR Extraction
Tesseract OCR
904
906
902
122
900

FIG. 9

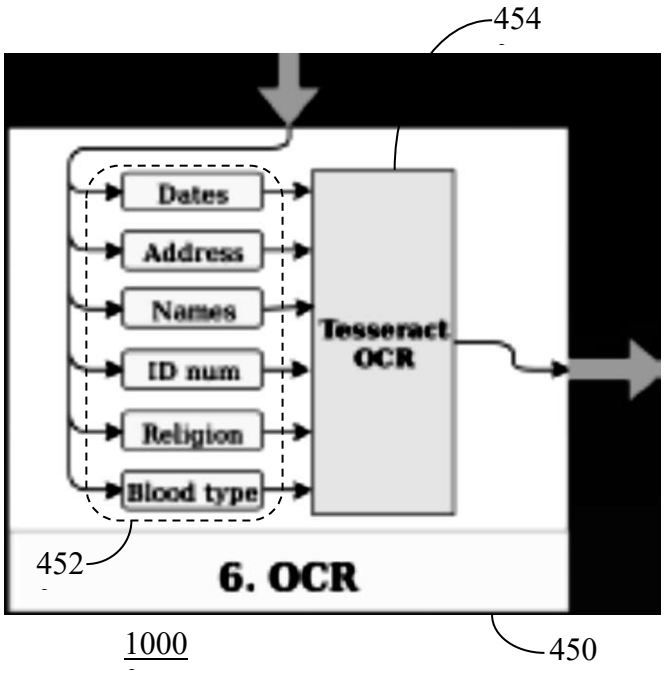


FIG. 10A

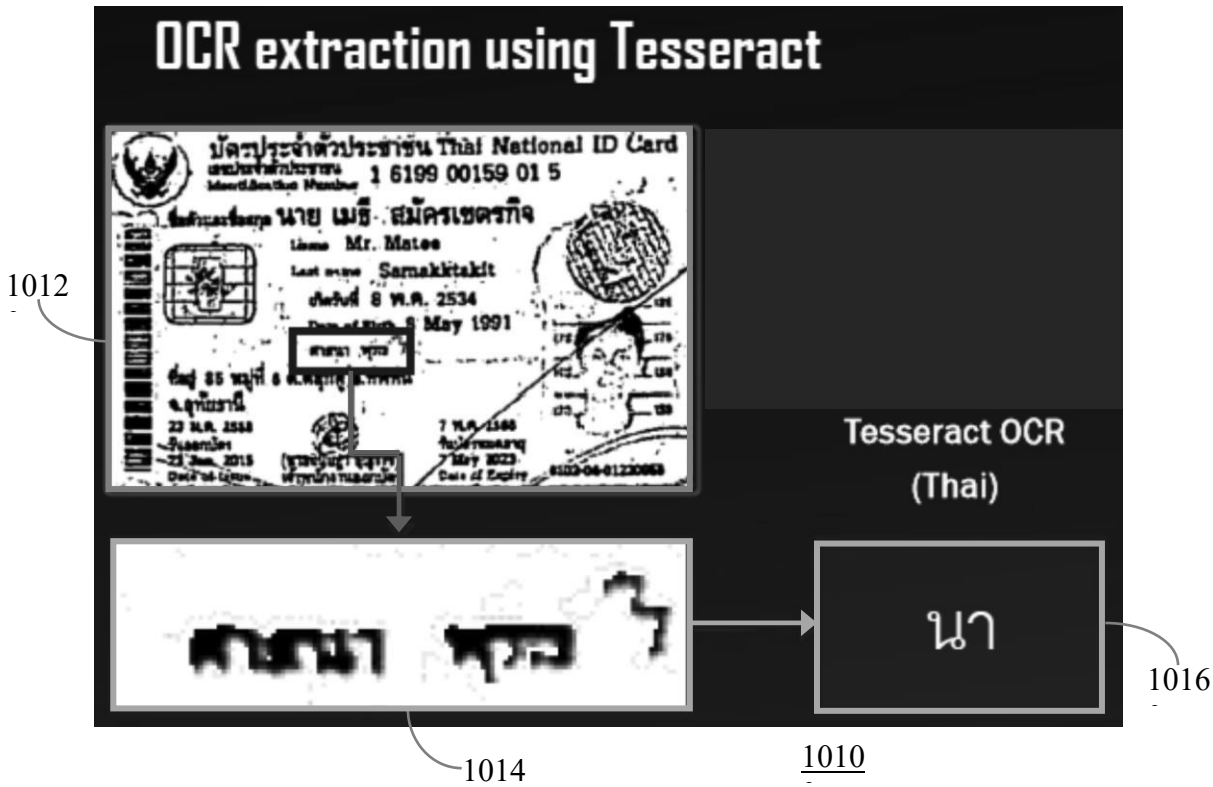


FIG. 10B



FIG. 10C

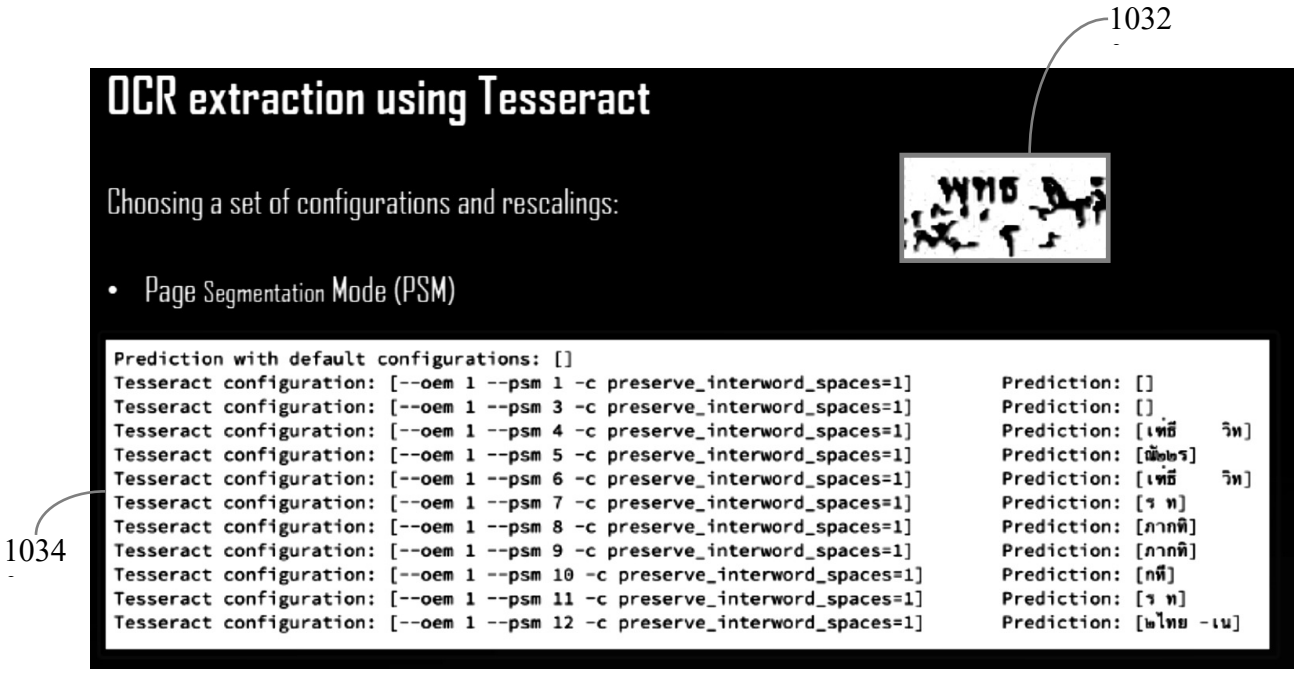


FIG. 10D

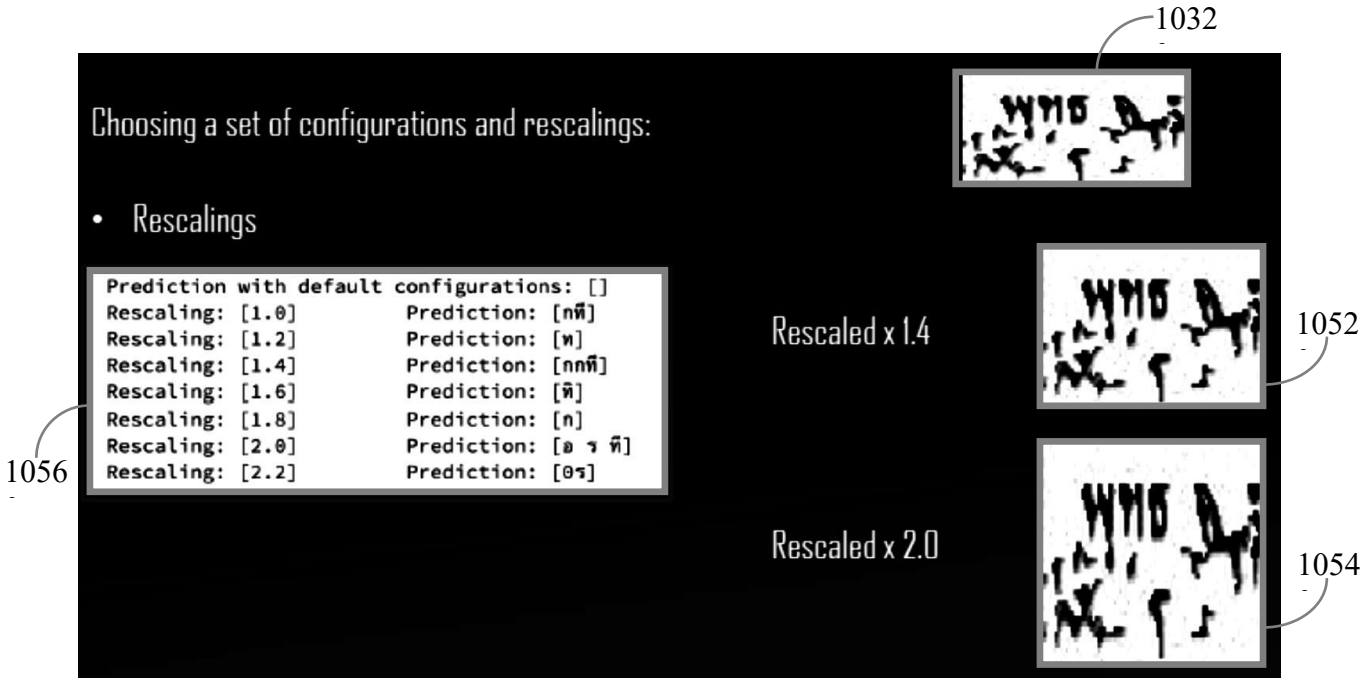


FIG. 10E

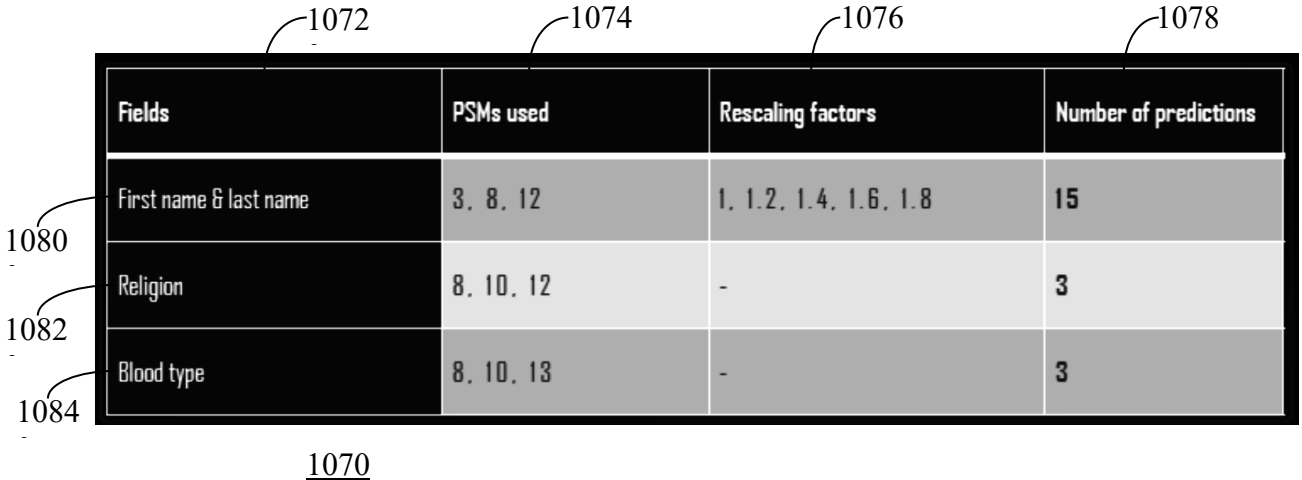
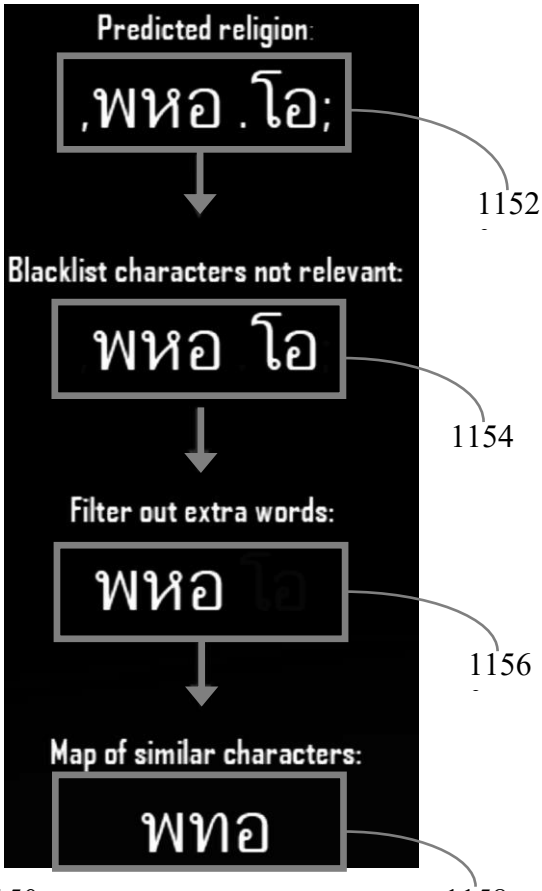


FIG. 10F



1100

FIG. 11A



1150

FIG. 11B

SPELLCHECKER

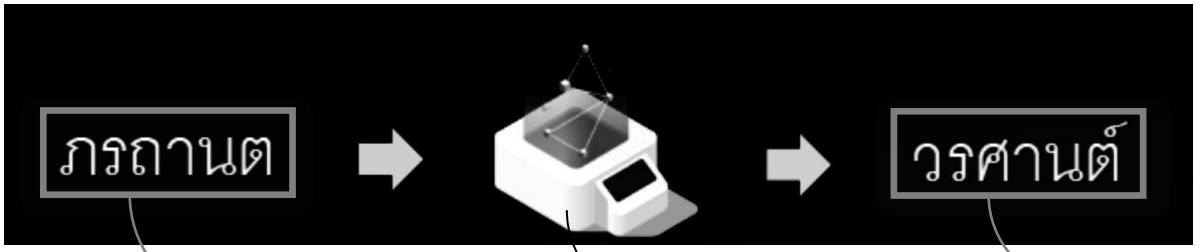


FIG. 12

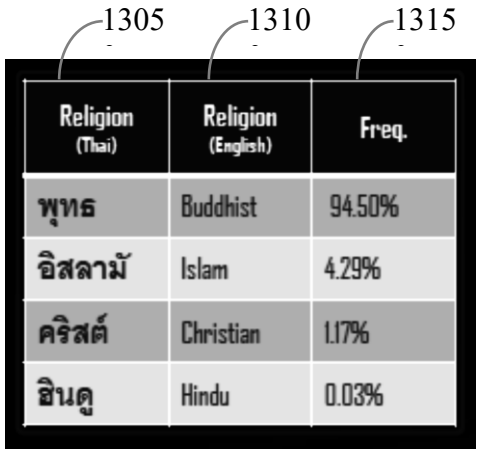
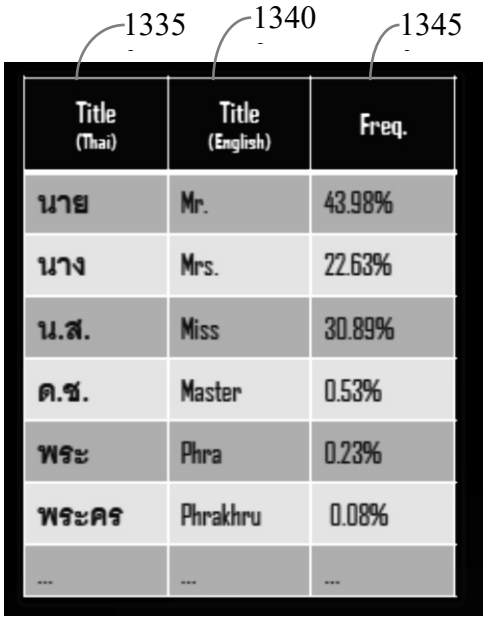


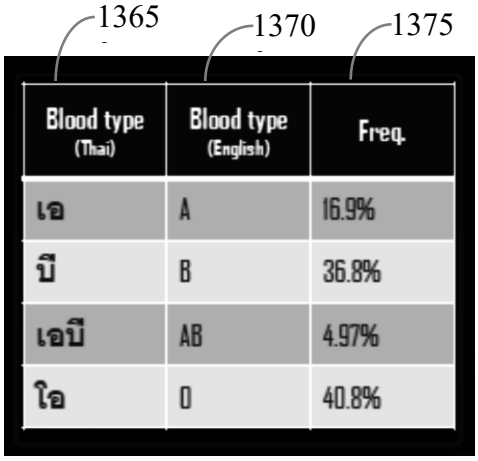
FIG. 13A



Title (Thai)	Title (English)	Freq.
นาย	Mr.	43.98%
นาง	Mrs.	22.63%
น.ส.	Miss	30.89%
ด.ช.	Master	0.53%
พระ	Phra	0.23%
พระคร	Phrakhru	0.08%
...	...	...

1330

FIG. 13B



Blood type (Thai)	Blood type (English)	Freq.
เอ	A	16.9%
บี	B	36.8%
เอบี	AB	4.97%
โอ	O	40.8%

FIG. 13C 1360

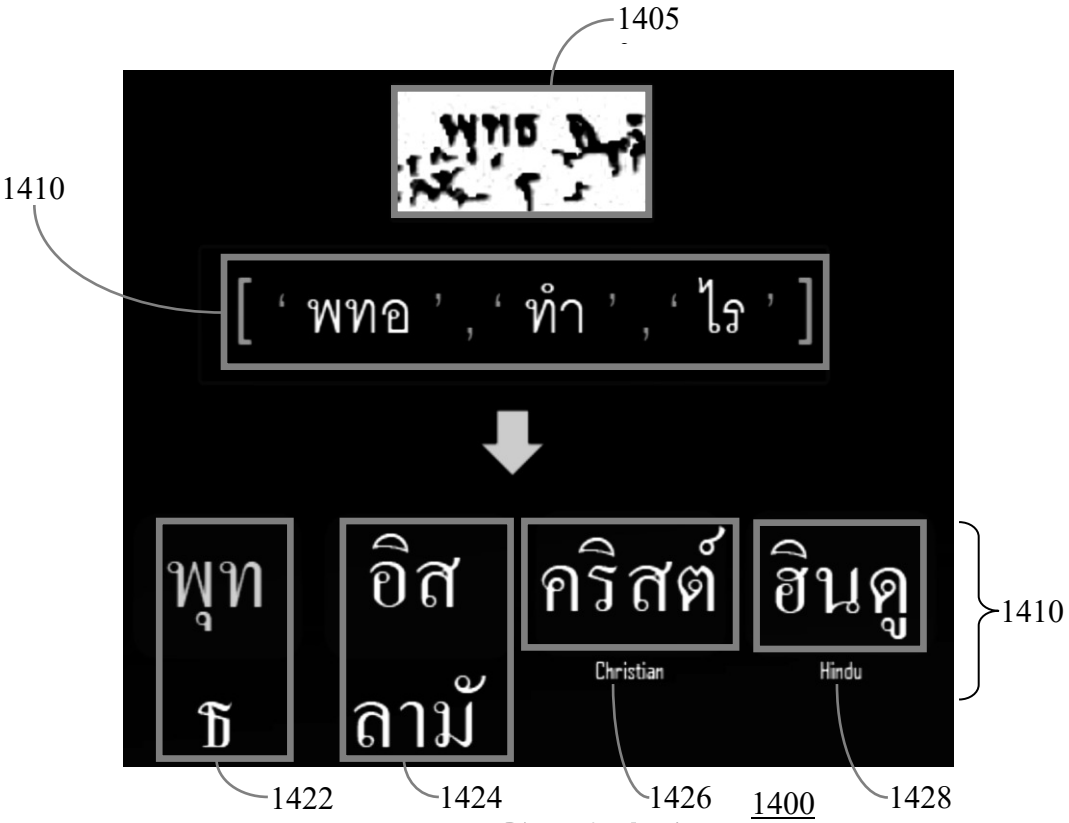


FIG. 14A

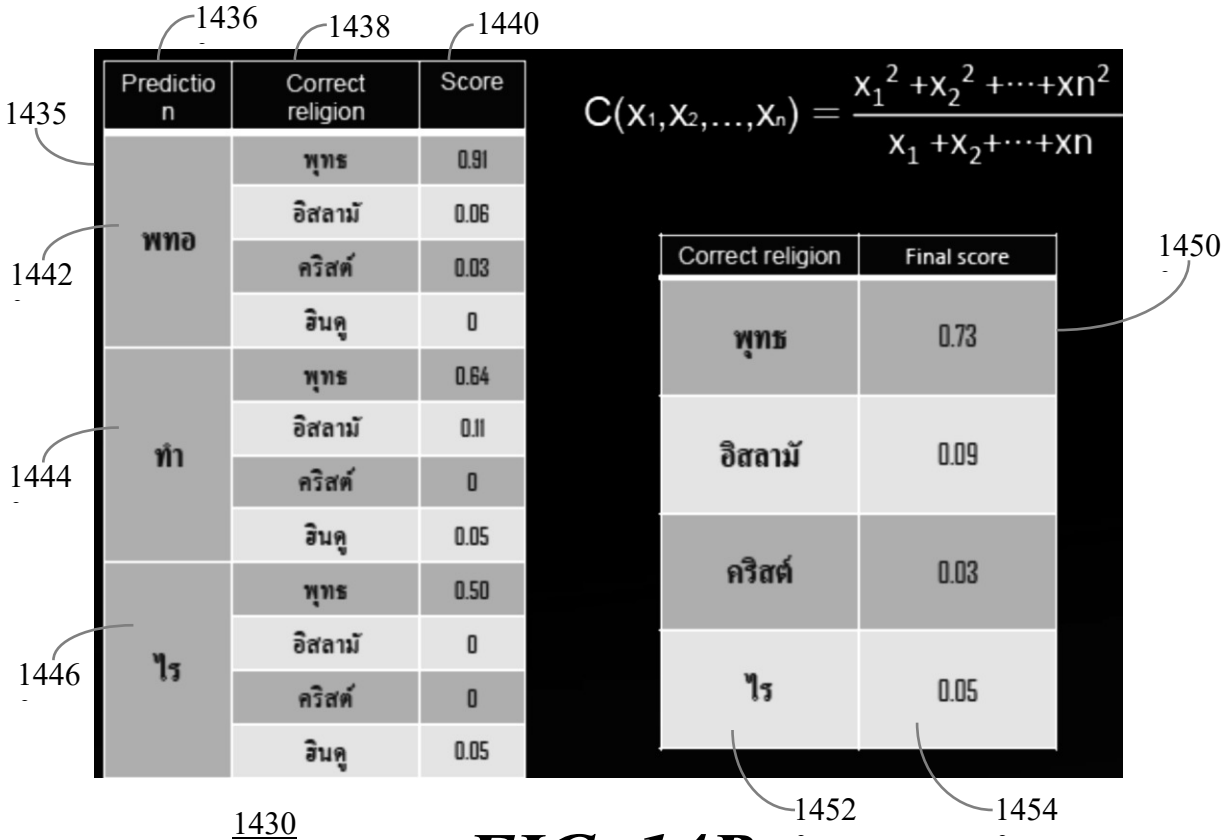
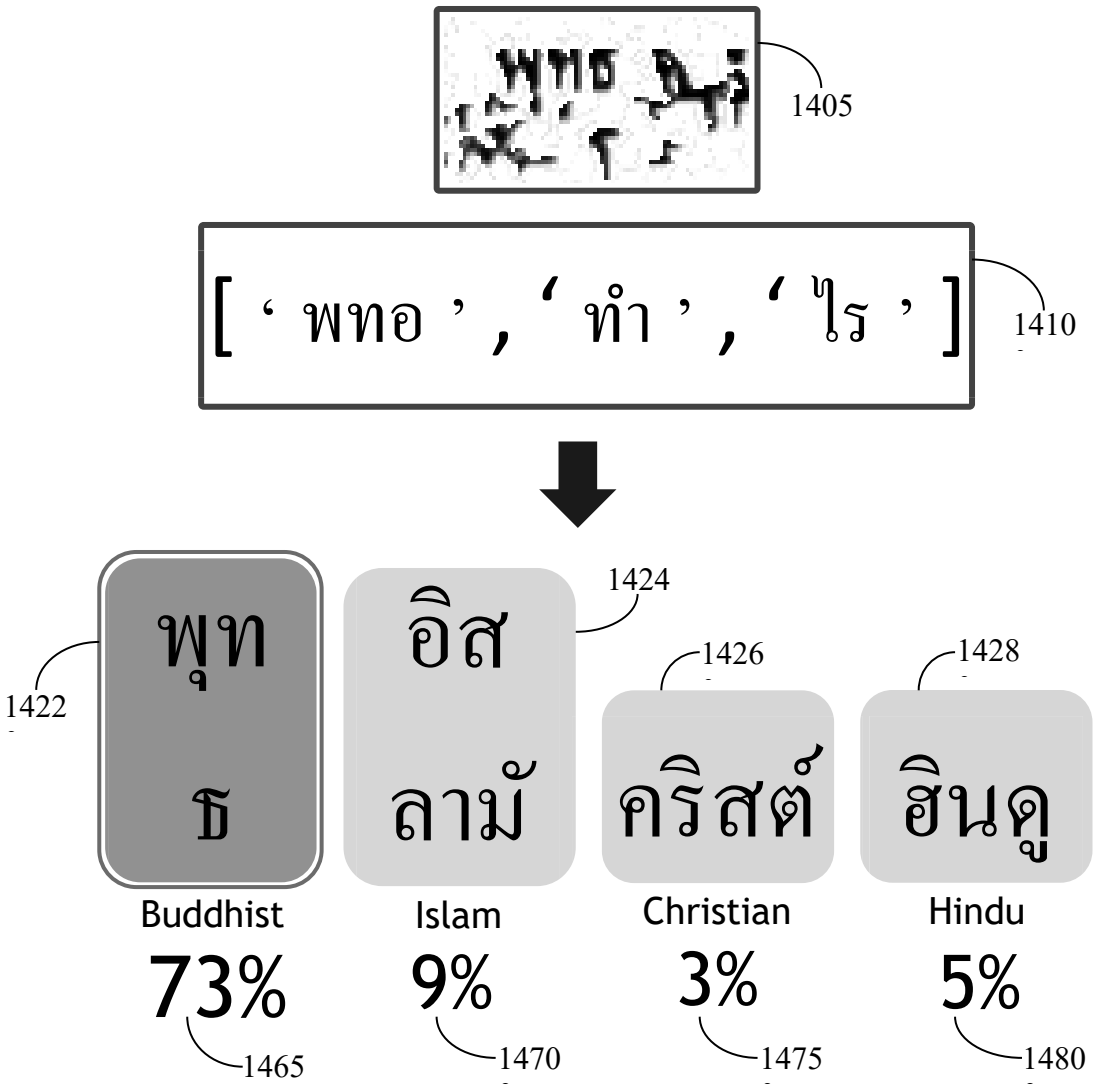


FIG. 14B


1405
[‘ พทอ ’ , ‘ ทำ ’ , ‘ ไร ’]
1410
1422
พุท
ธ
1424
อิส
ลามั
1426
คริสต์
1428
ฮินดู
Buddhist
73%
1465
Islam
9%
1470
Christian
3%
1475
Hindu
5%
1480

1460 ***FIG. 14C***

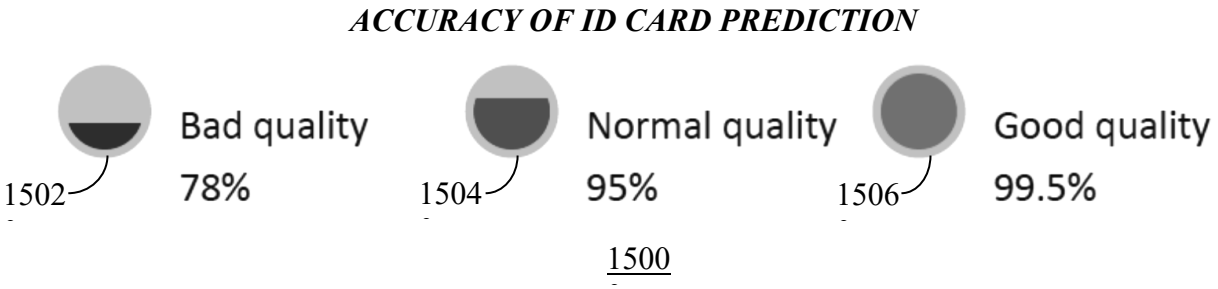


FIG. 15A

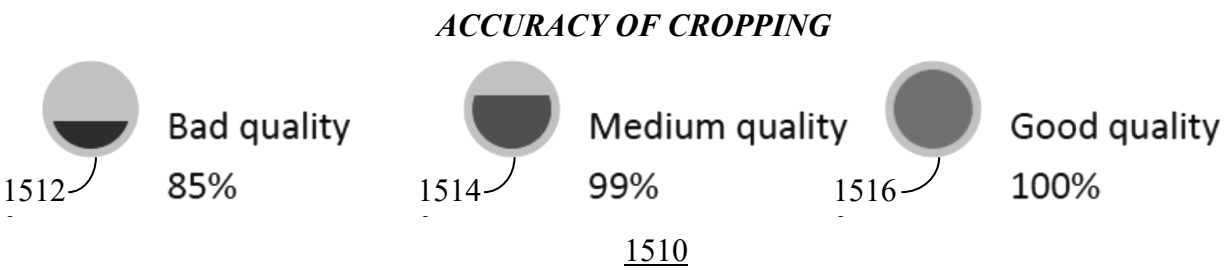


FIG. 15B

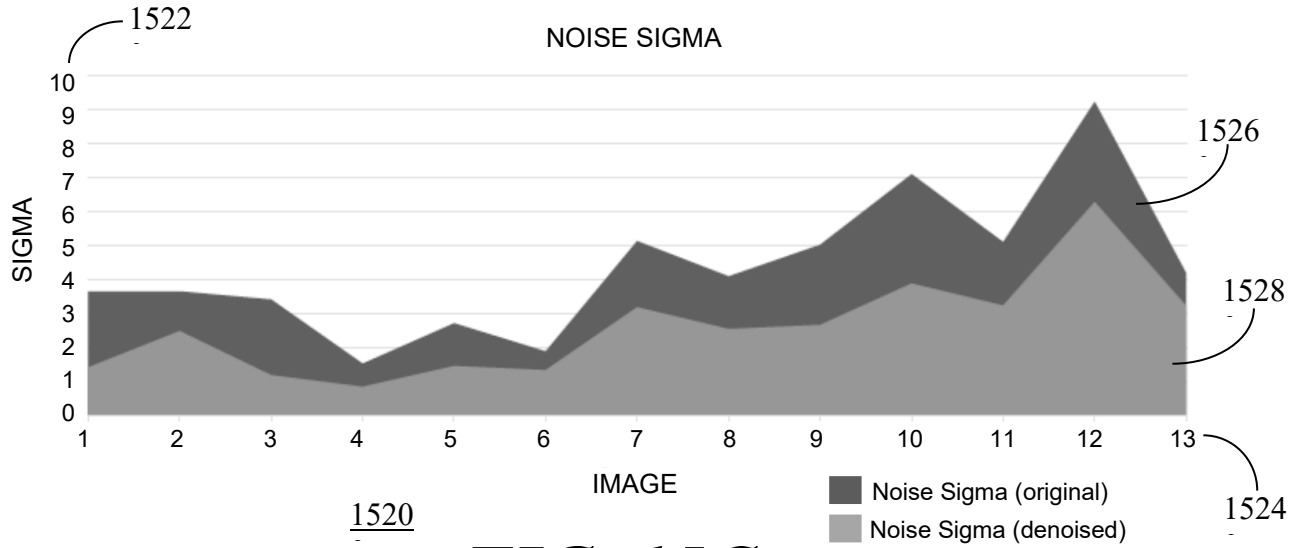
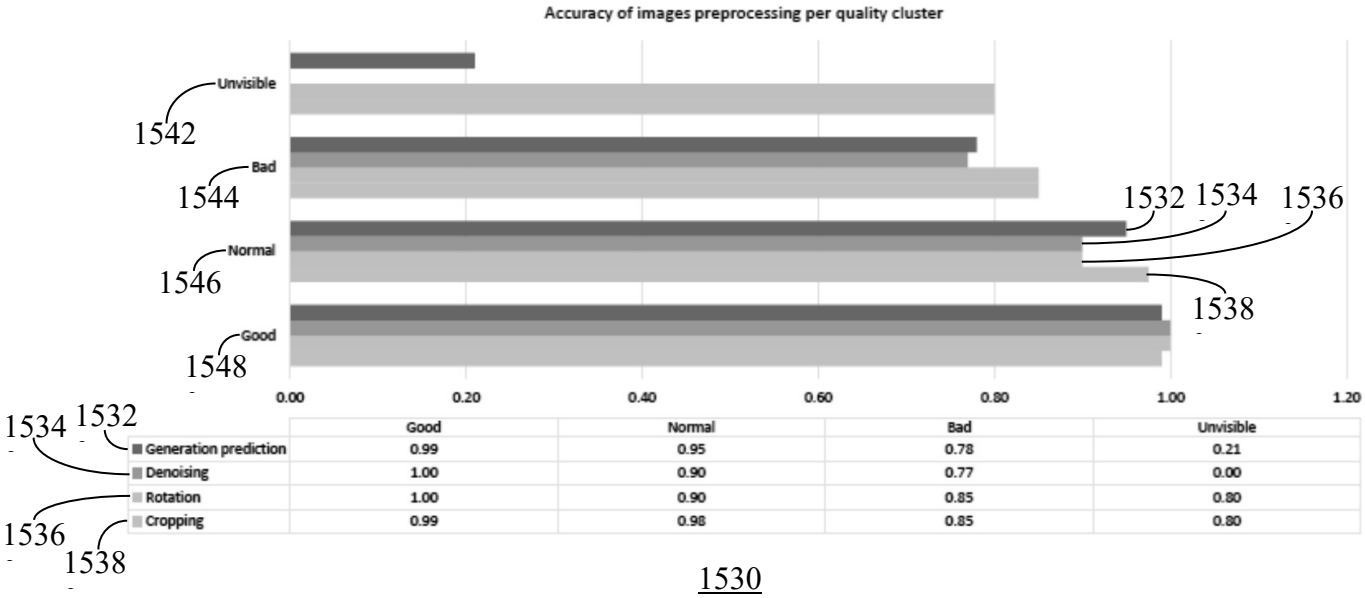
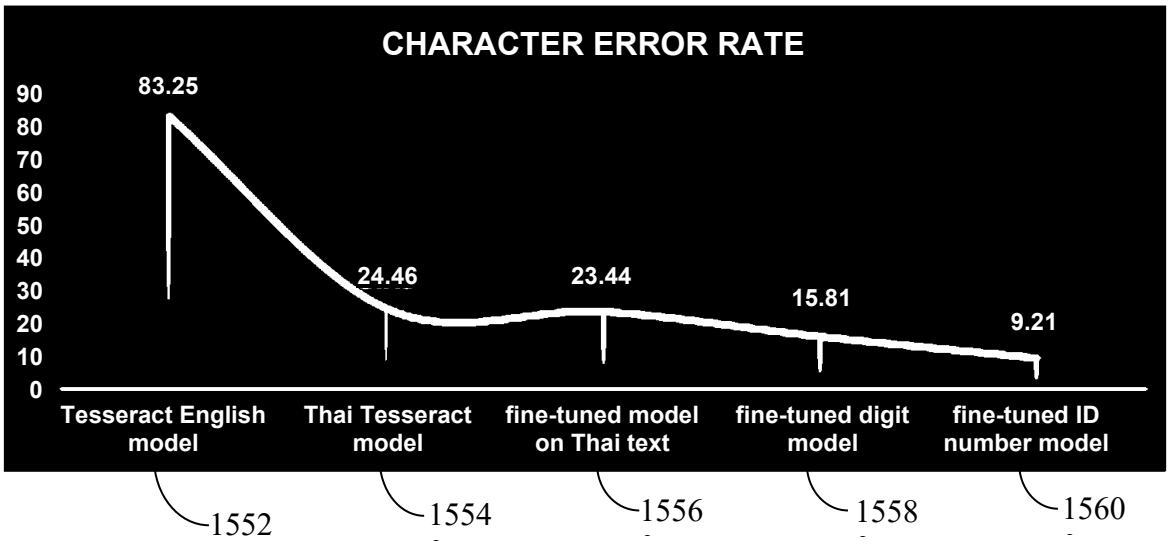


FIG. 15C



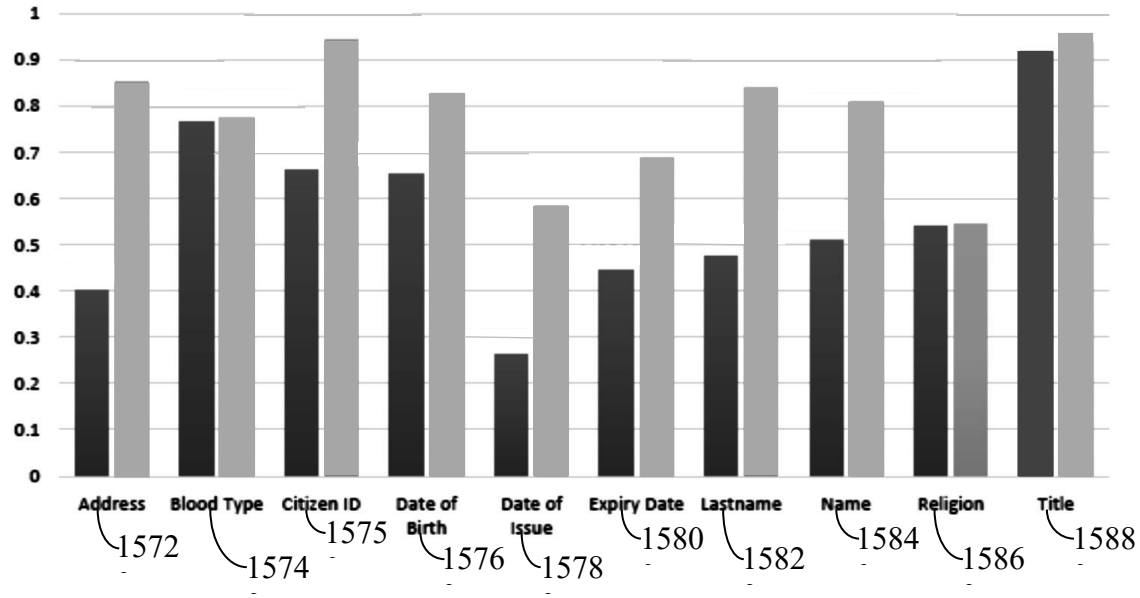
1530

FIG. 15D



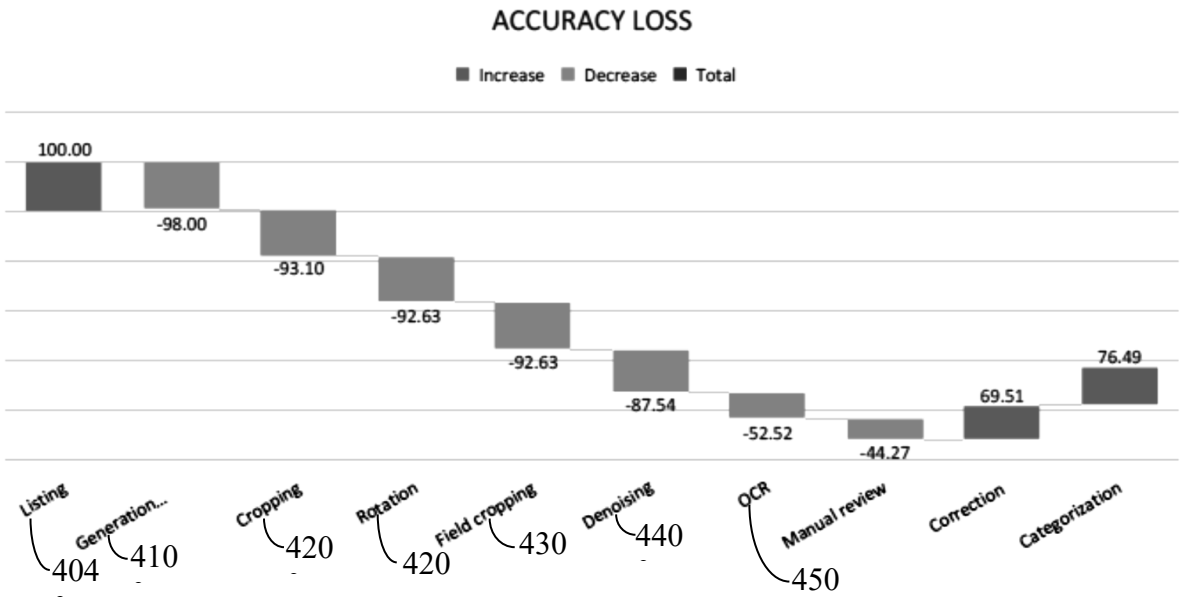
1550

FIG. 15E



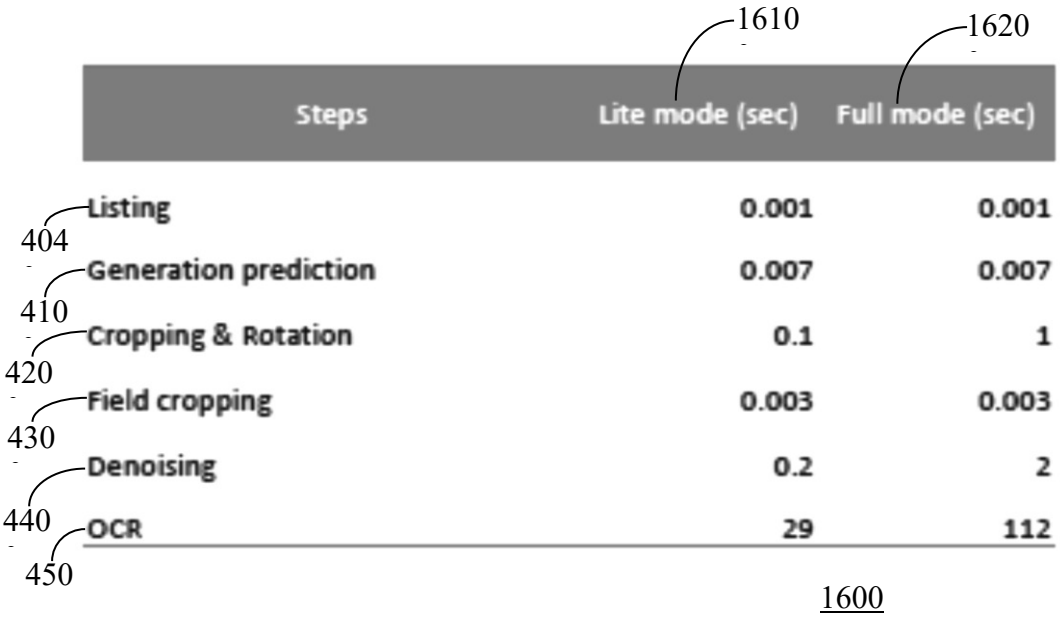
1570

FIG. 15F



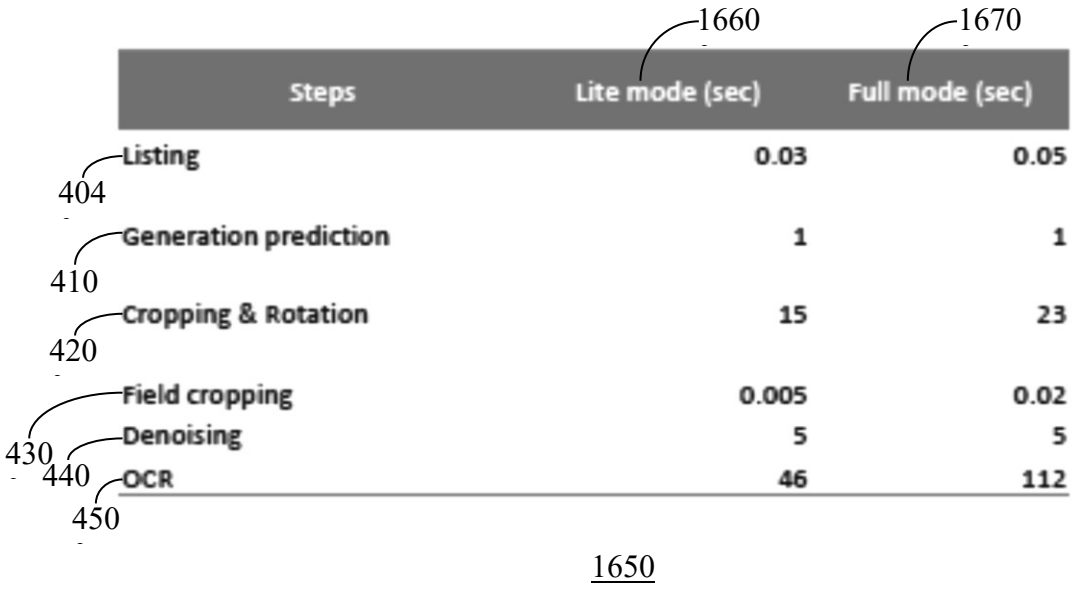
1590

FIG. 15G



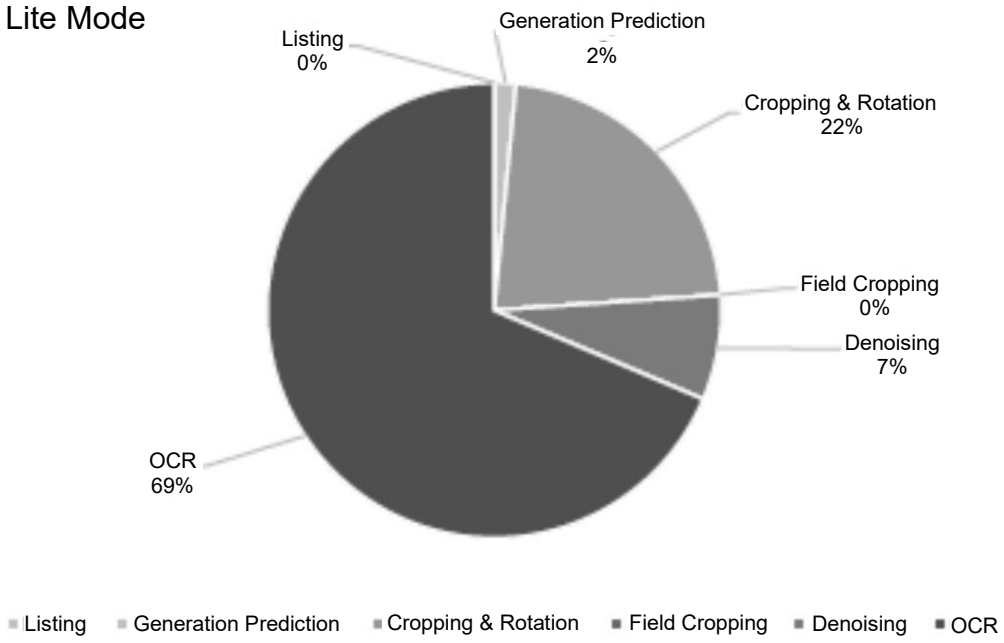
Steps	Lite mode (sec)	Full mode (sec)
Listing	0.001	0.001
Generation prediction	0.007	0.007
Cropping & Rotation	0.1	1
Field cropping	0.003	0.003
Denoising	0.2	2
OCR	29	112

FIG. 16A



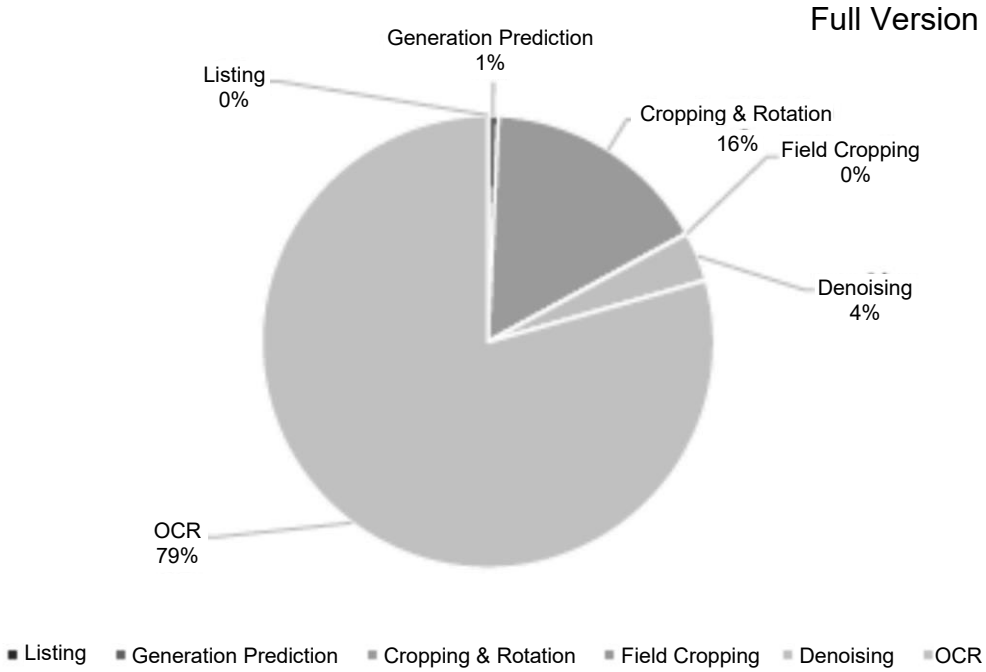
Steps	Lite mode (sec)	Full mode (sec)
Listing	0.03	0.05
Generation prediction	1	1
Cropping & Rotation	15	23
Field cropping	0.005	0.02
Denoising	5	5
OCR	46	112

FIG. 16B



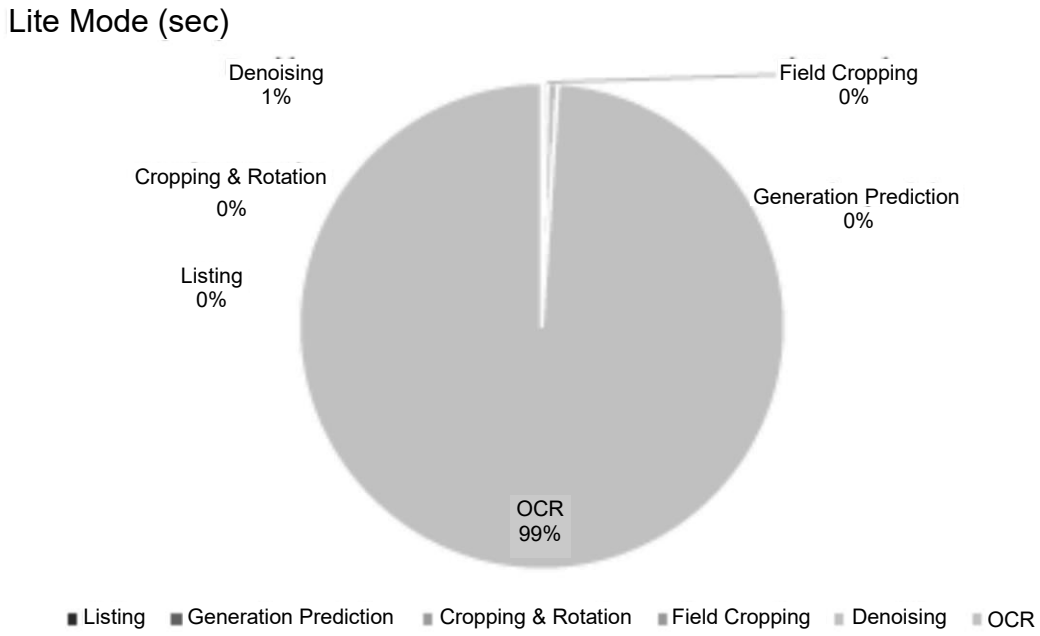
1700

FIG. 17A



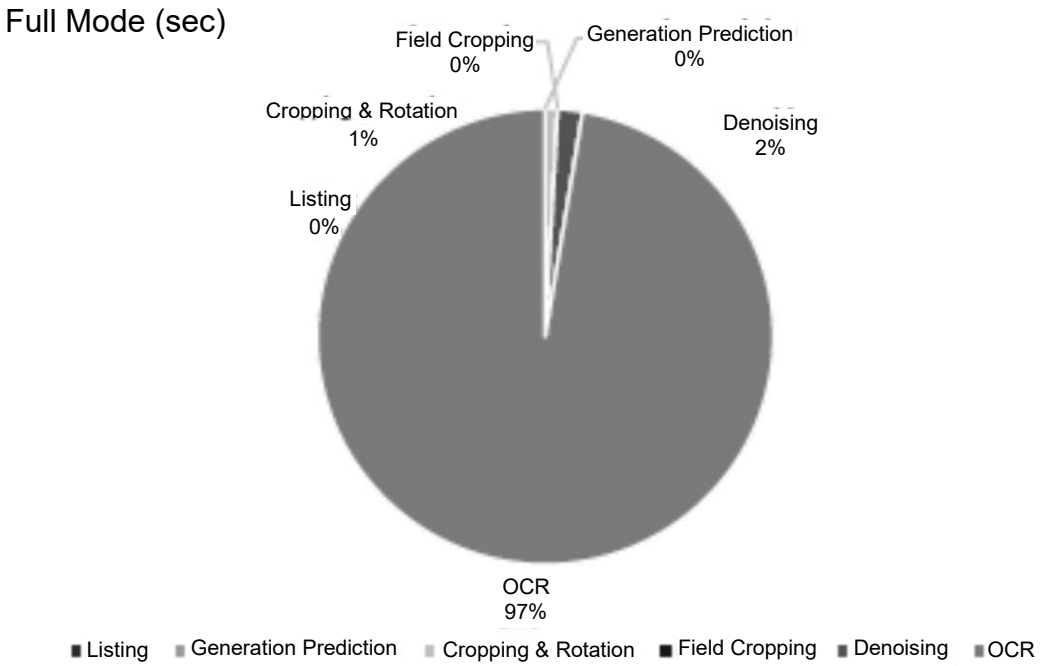
1750

FIG. 17B



1800

FIG. 18A



1850

FIG. 18B