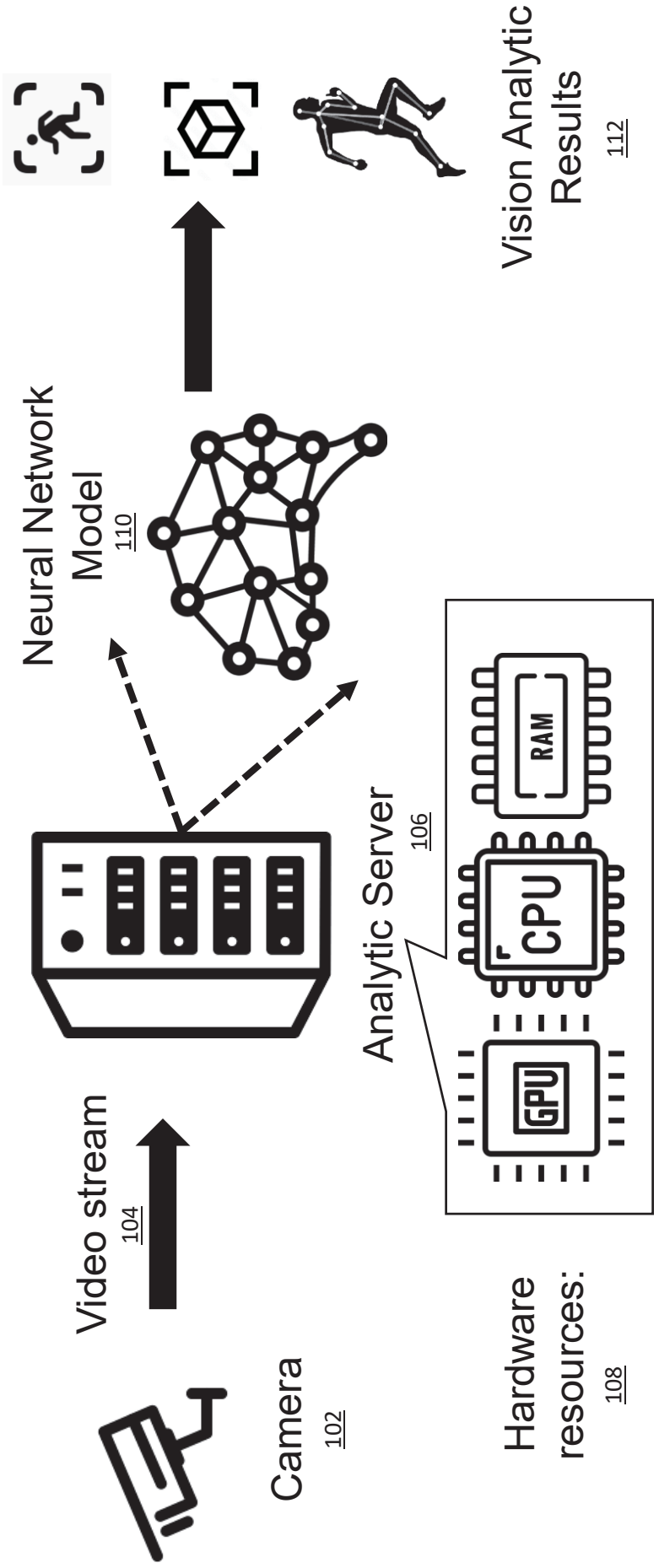
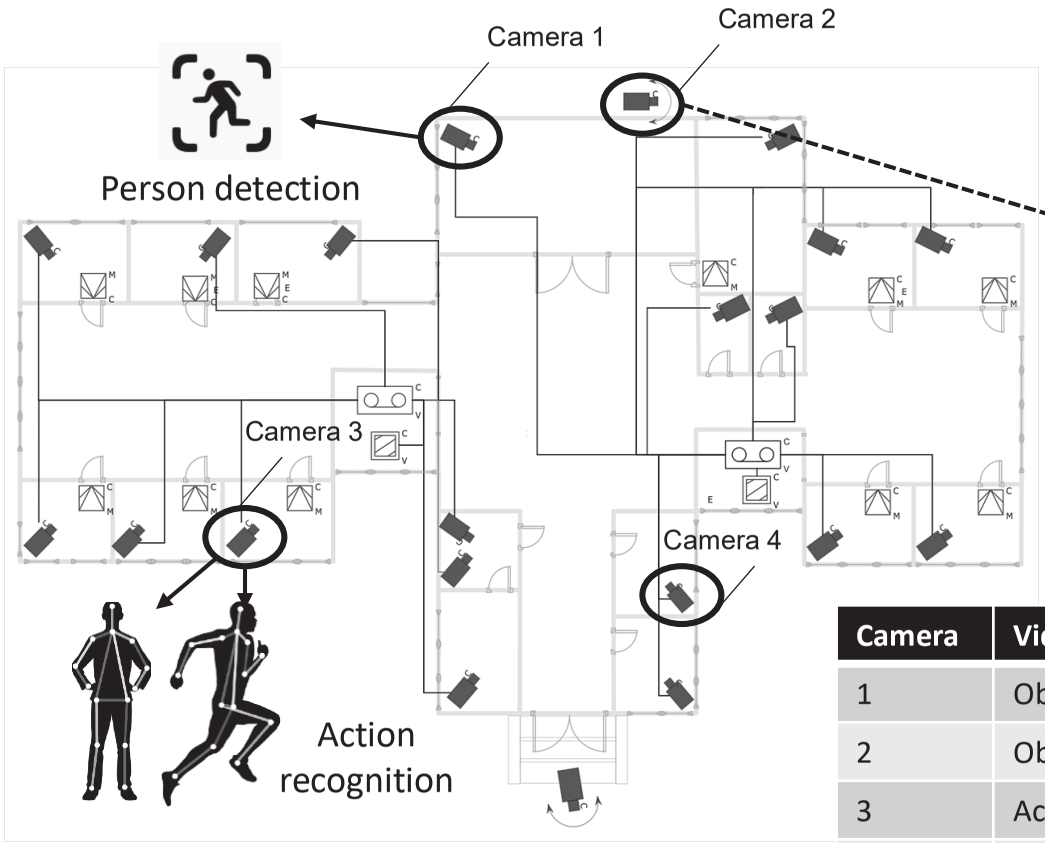

Vision Analytic Results
112
Neural Network Model
110
Video stream
104
Analytic Server
106
RAM
CPU
GPU
Hardware resources:
108
Camera
102
100

Figure 1



Camera	Video Analysis Task	Target Classes
1	Object classification	Person
2	Object classification	Person, box
3	Action recognition	Walk, fight, talk
4	Action recognition	Walk, talk
…	…	…
N	…	…

Figure 2

Video Stream	Application Requirement
1	Person, vehicle, bike, box
2	Person, box
3	Person, vehicle
4	Person
...	
N	Bike

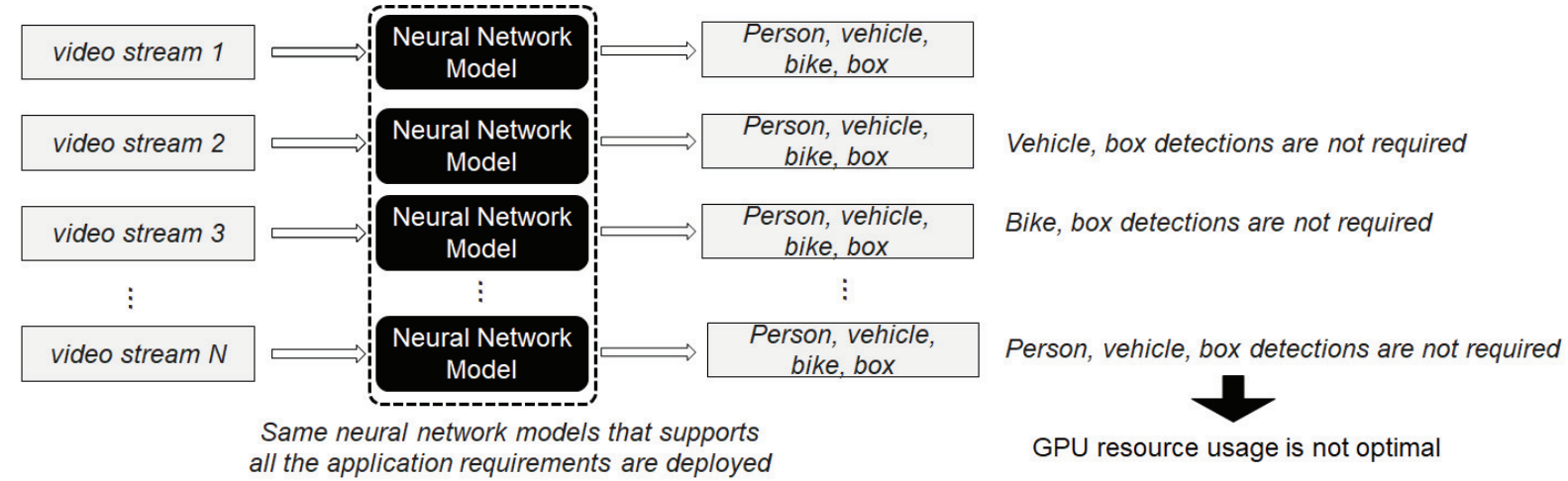


Figure 3

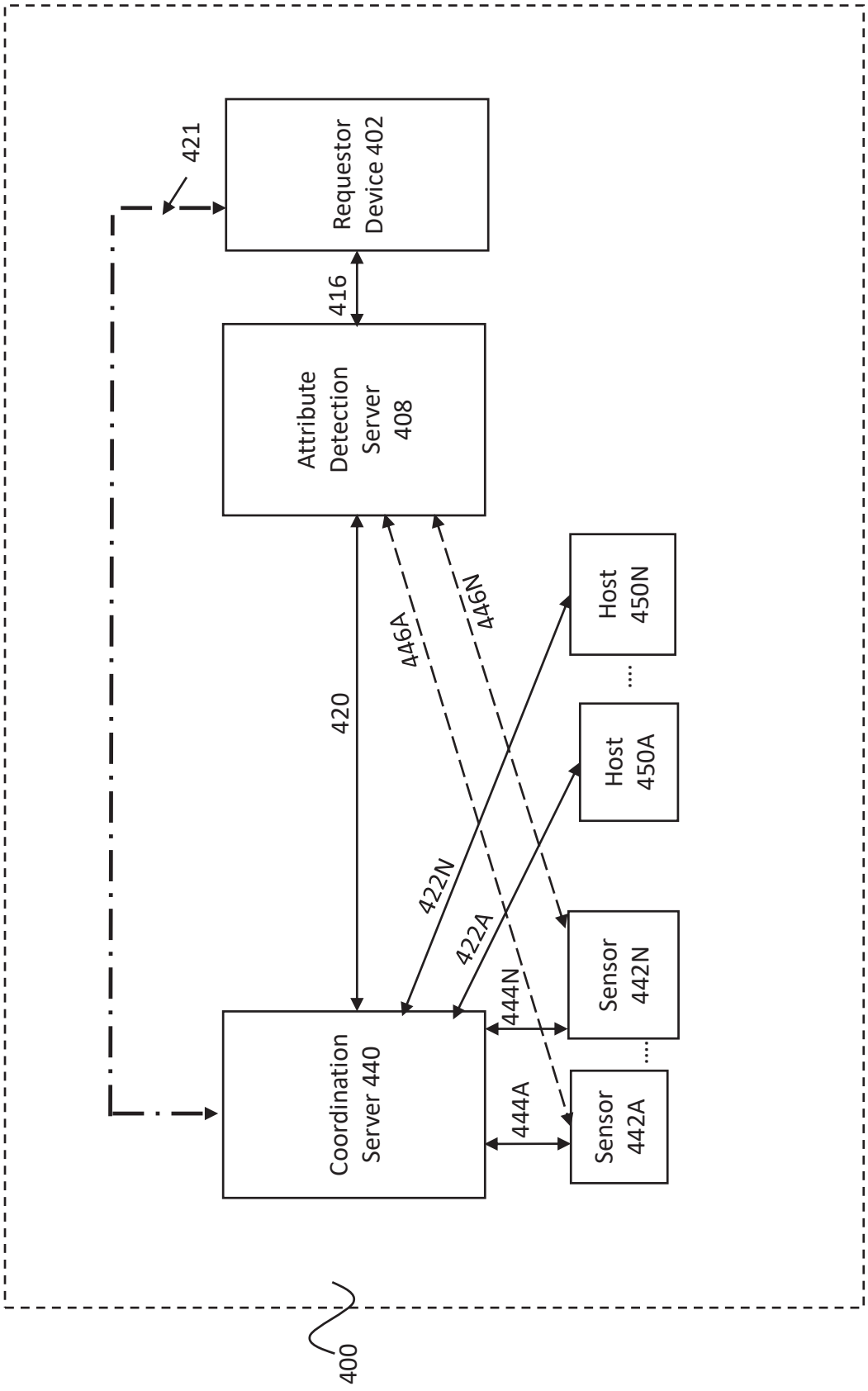

400
421
Requestor Device 402
416
Attribute Detection Server 408
420
446A
446N
422N
422A
444N
444A
Coordination Server 440
Host 450N
Host 450A
Sensor 442N
Sensor 442A

Figure 4

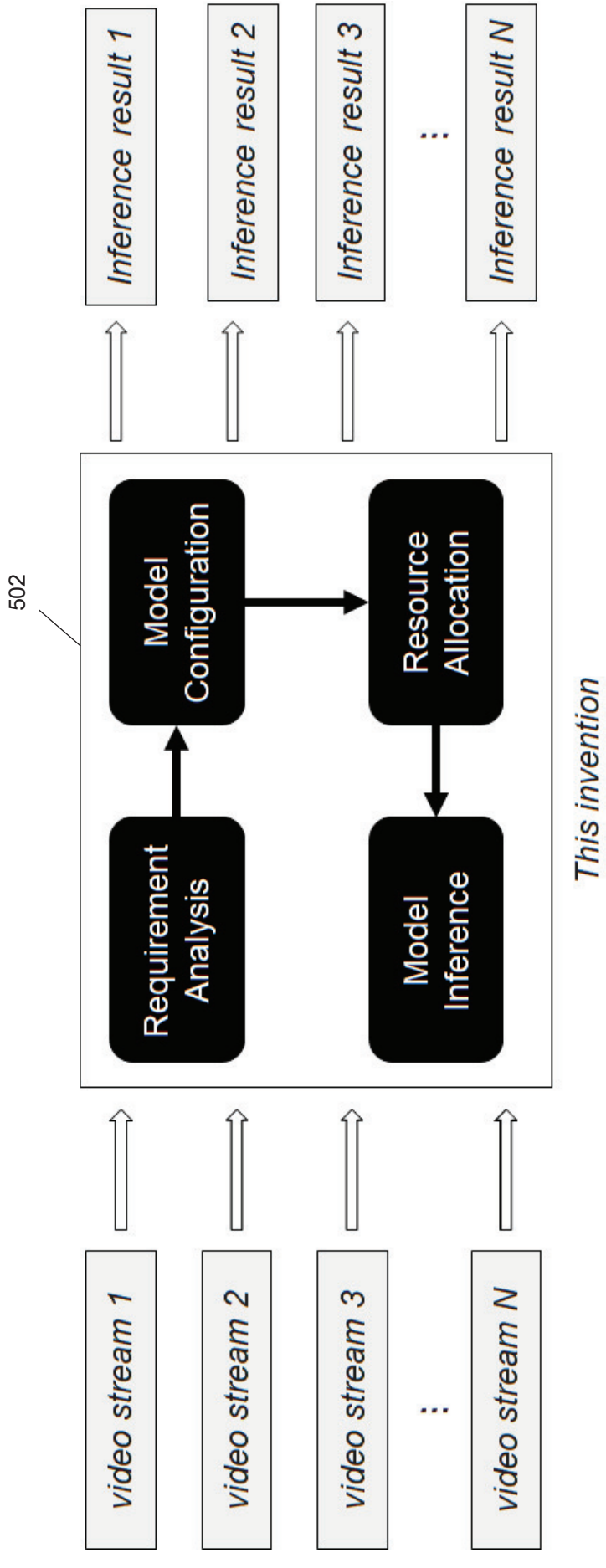

video stream 1
video stream 2
video stream 3
video stream N
500
502
Requirement Analysis
Model Configuration
Model Inference
Resource Allocation
This invention
Inference result 1
Inference result 2
Inference result 3
Inference result N

Figure 5

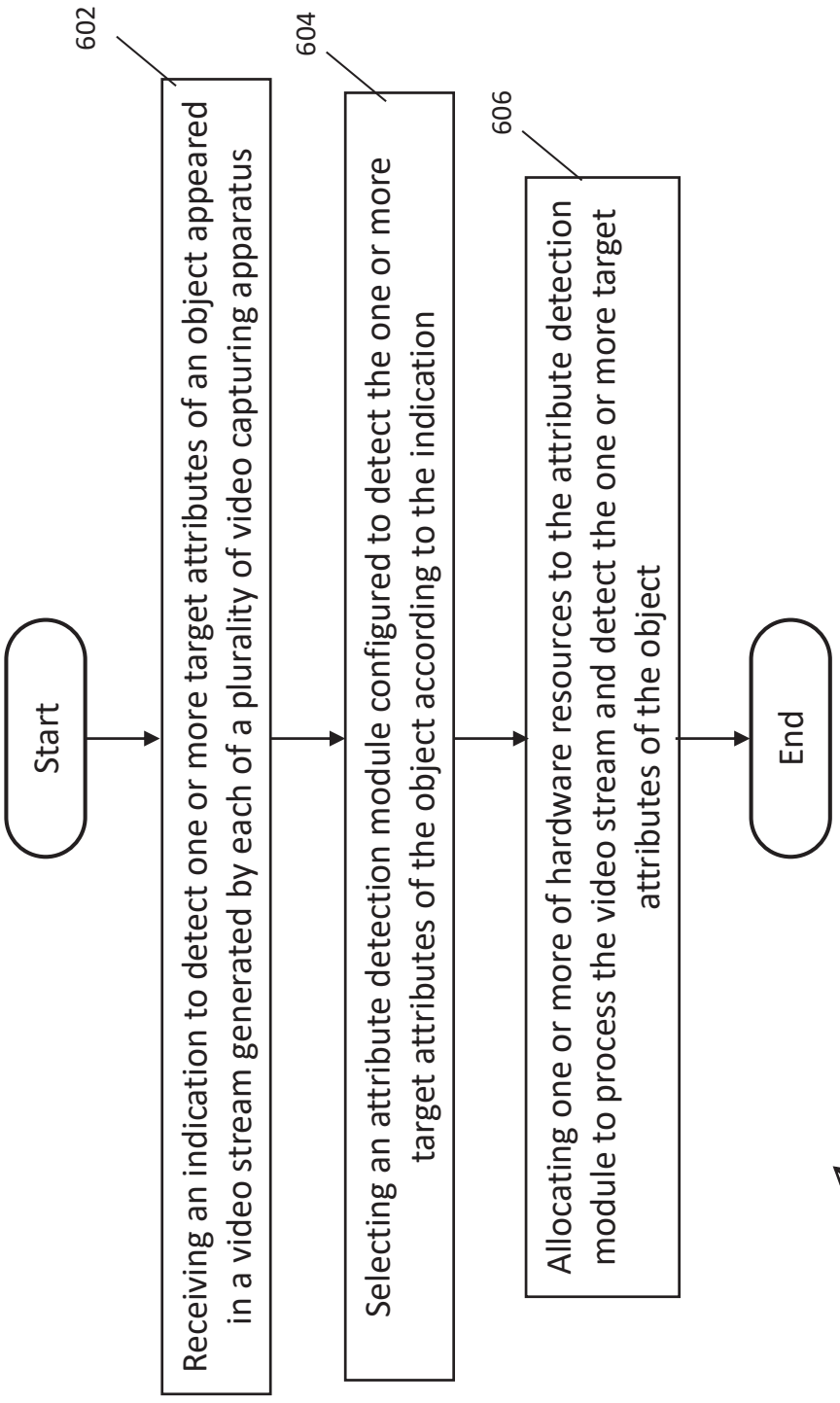

Start
602
Receiving an indication to detect one or more target attributes of an object appeared in a video stream generated by each of a plurality of video capturing apparatus
604
Selecting an attribute detection module configured to detect the one or more target attributes of the object according to the indication
606
Allocating one or more of hardware resources to the attribute detection module to process the video stream and detect the one or more target attributes of the object
End
600

Figure 6

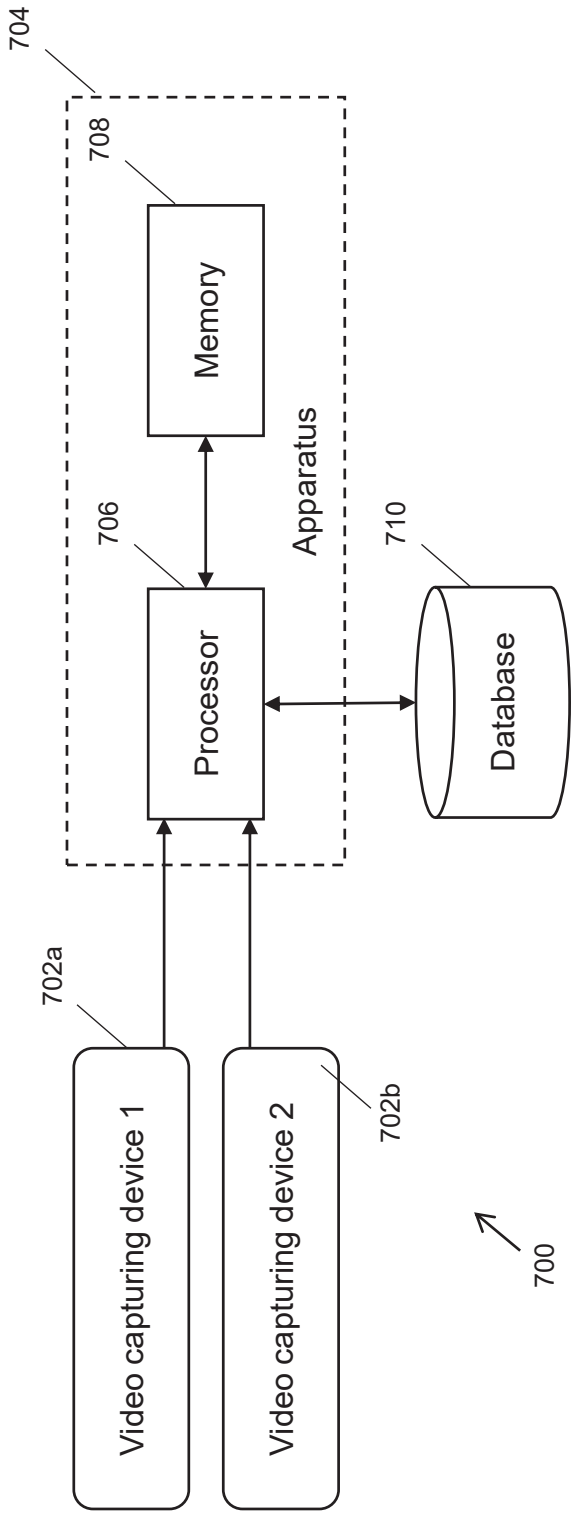

704
708
Memory
Apparatus
706
710
Processor
Database
702a
Video capturing device 1
Video capturing device 2
702b
700

Figure 7

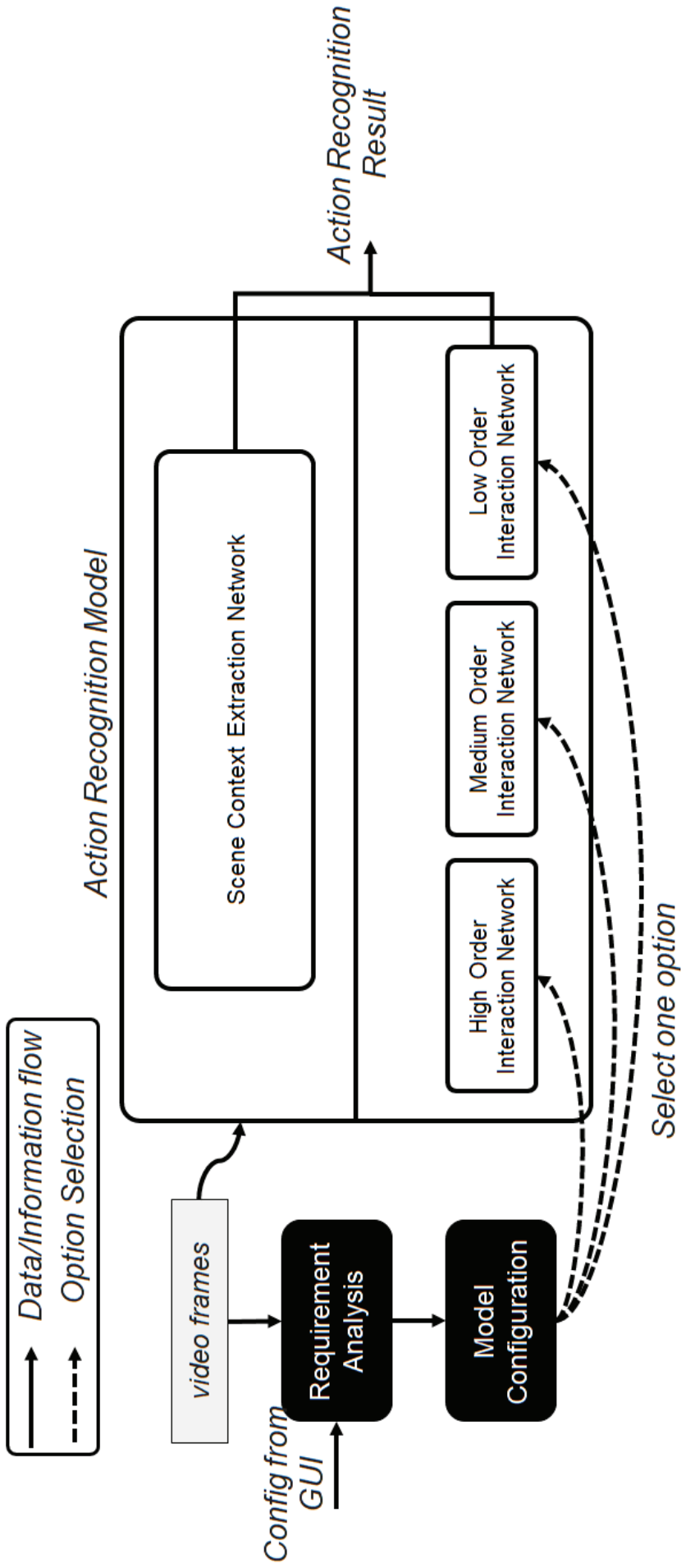

Data/Information flow
Option Selection
Action Recognition Model
Scene Context Extraction Network
High Order Interaction Network
Medium Order Interaction Network
Low Order Interaction Network
Action Recognition Result
video frames
Requirement Analysis
Model Configuration
Config from GUI
Select one option
800

Figure 8

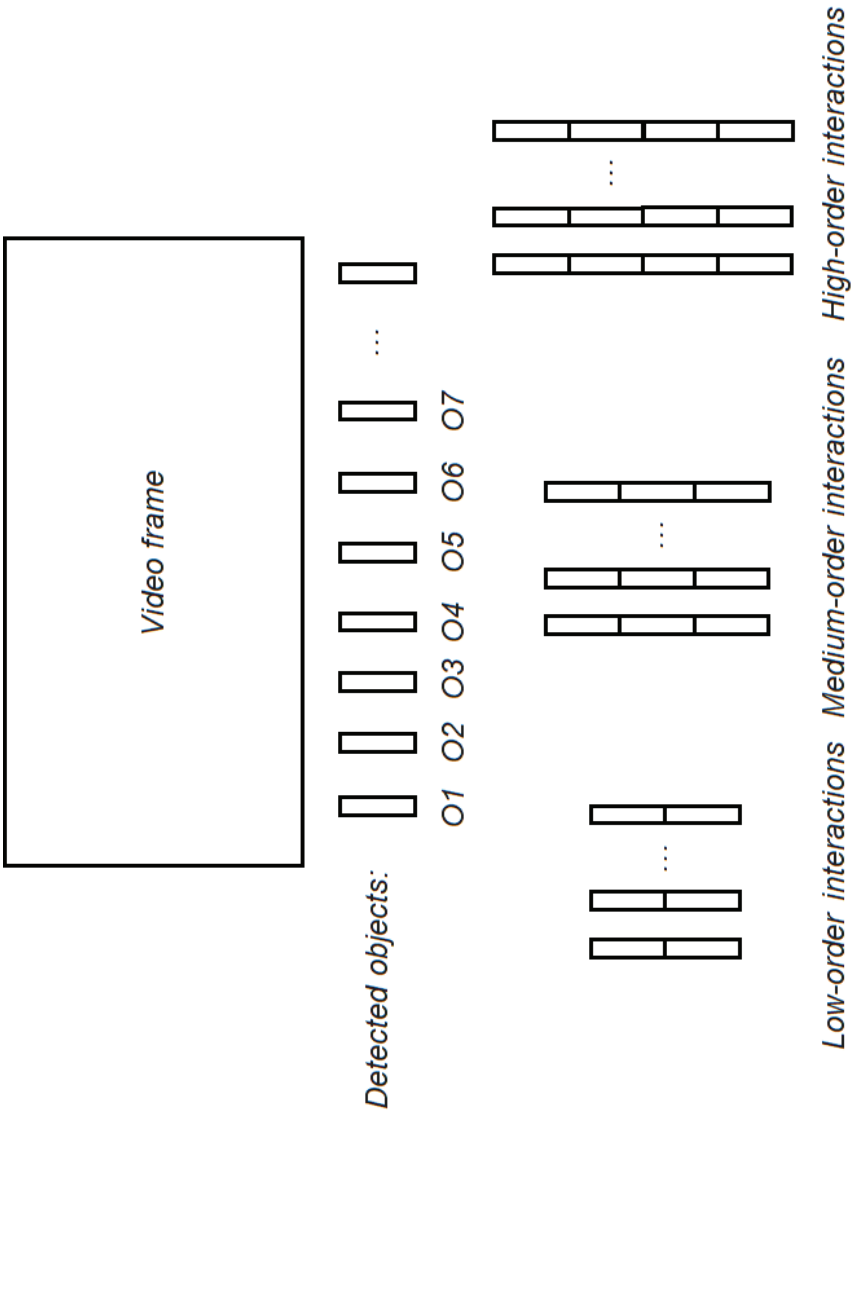

Video frame
Detected objects:
O1
O2
O3
O4
O5
O6
O7
Low-order interactions
Medium-order interactions
High-order interactions
900

Figure 9

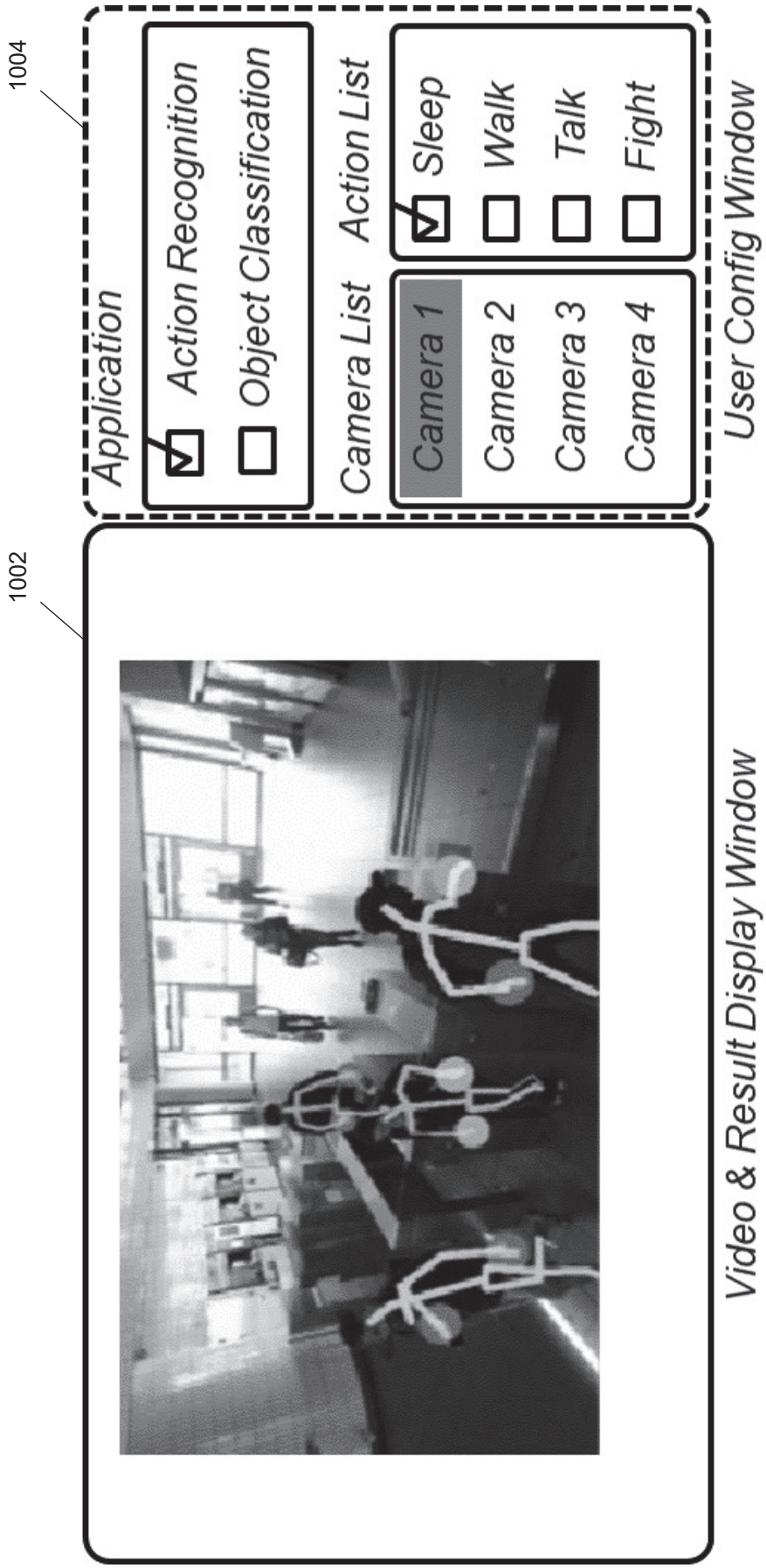

1000
1002
1004
Video & Result Display Window
User Config Window
Application
Action Recognition
Object Classification
Camera List
Camera 1
Camera 2
Camera 3
Camera 4
Action List
Sleep
Walk
Talk
Fight

Figure 10

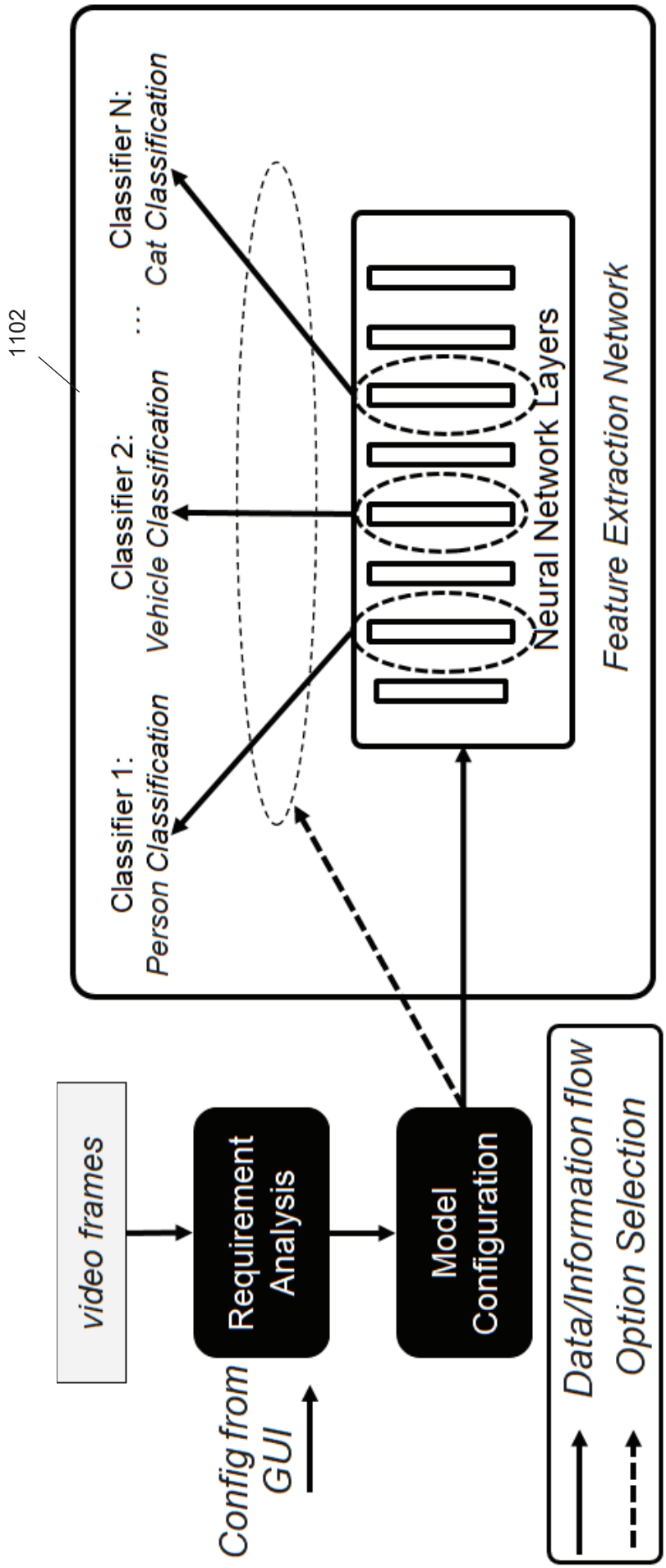

1102
Classifier 1:
Person Classification
Classifier 2:
Vehicle Classification
...
Classifier N:
Cat Classification
Neural Network Layers
Feature Extraction Network
video frames
Requirement Analysis
Model Configuration
Config from GUI
Data/Information flow
Option Selection
1100

Figure 11

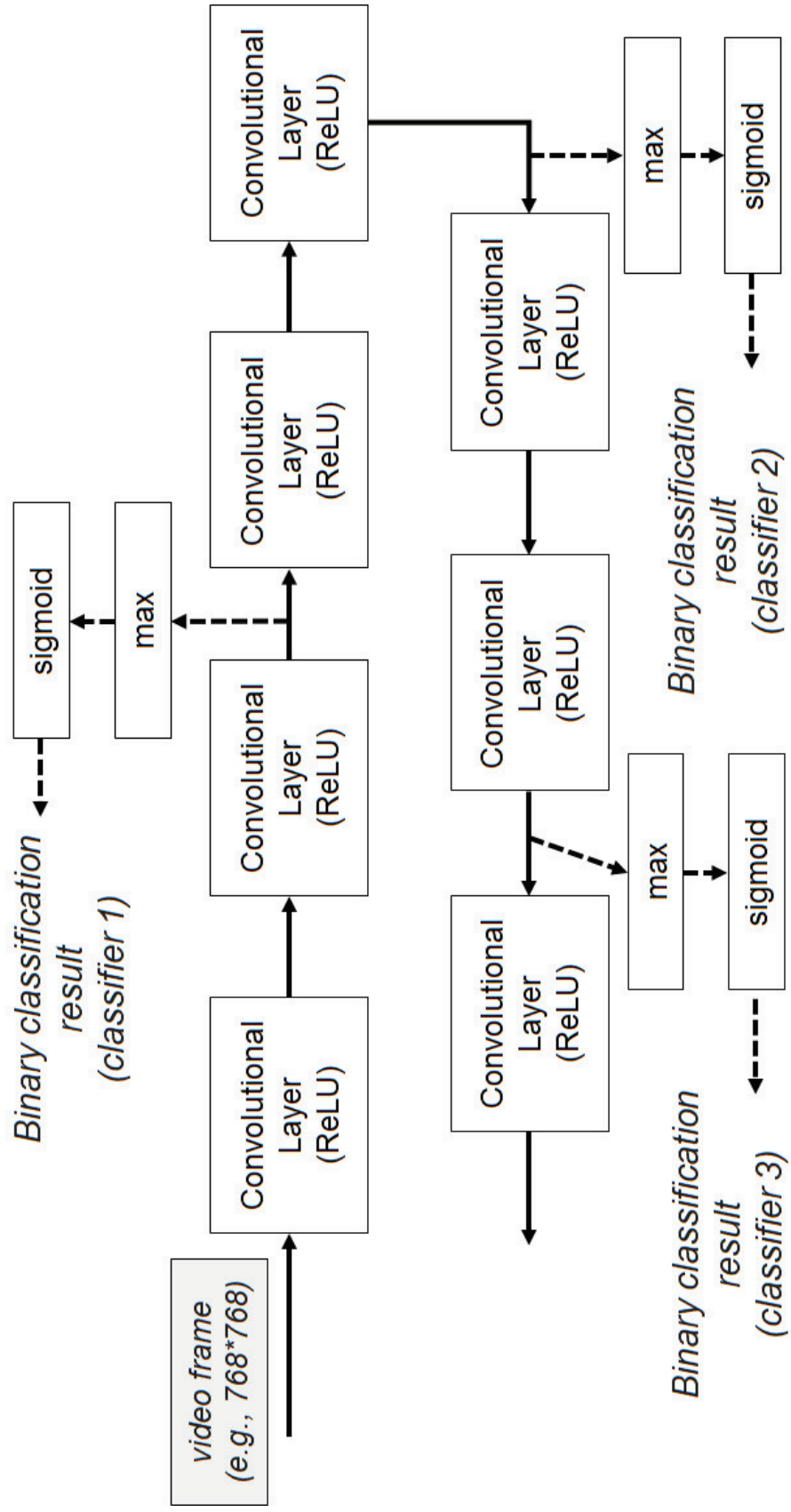

video frame
(e.g., 768*768)
Convolutional Layer (ReLU)
Convolutional Layer (ReLU)
Convolutional Layer (ReLU)
Convolutional Layer (ReLU)
max
sigmoid
Binary classification result (classifier 1)
Convolutional Layer (ReLU)
Convolutional Layer (ReLU)
Convolutional Layer (ReLU)
max
sigmoid
Binary classification result (classifier 2)
Convolutional Layer (ReLU)
max
sigmoid
Binary classification result (classifier 3)
1200

Figure 12

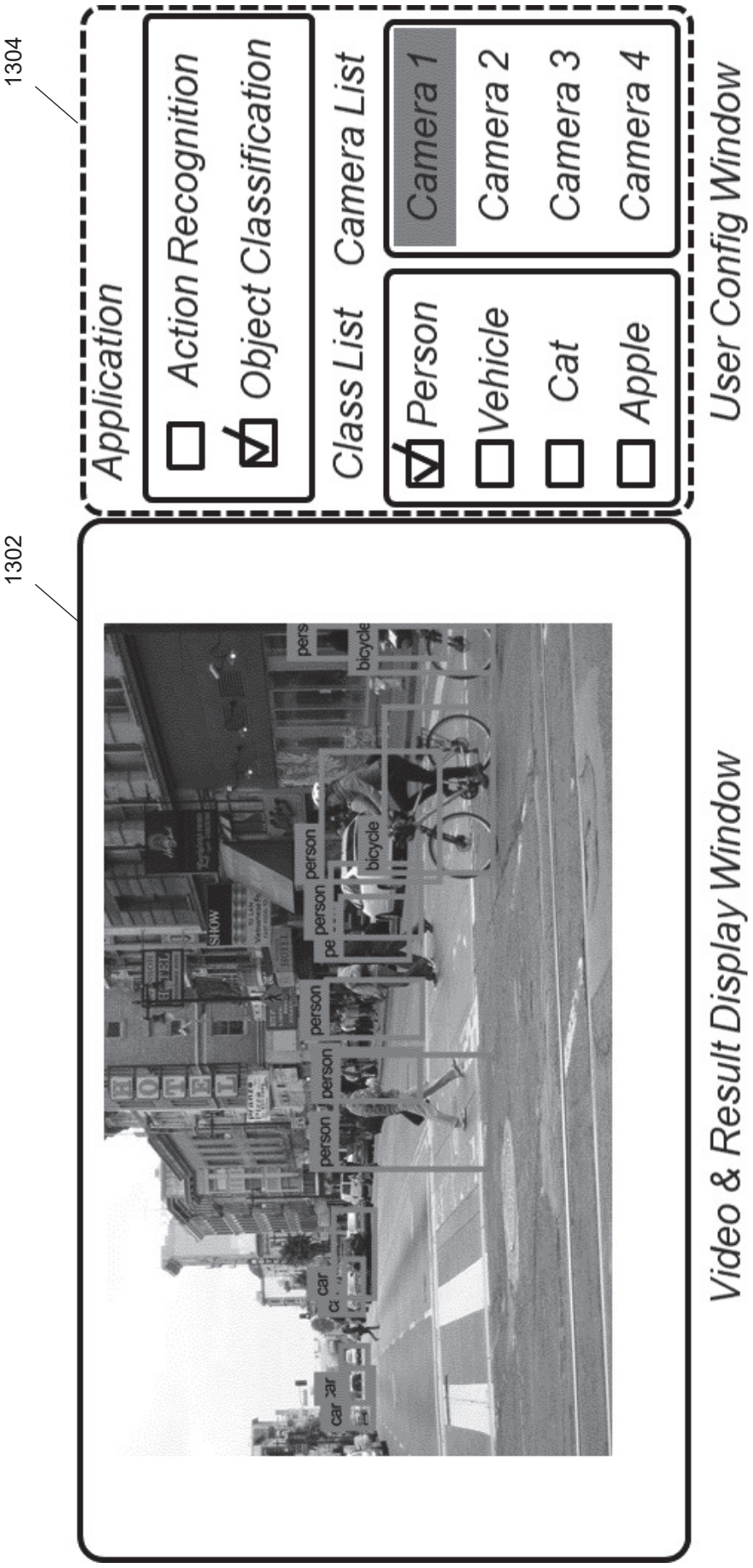

1304
1302
1300
Application
Action Recognition
Object Classification
Class List
Camera List
Person
Vehicle
Cat
Apple
Camera 1
Camera 2
Camera 3
Camera 4
User Config Window
Video & Result Display Window

Figure 13

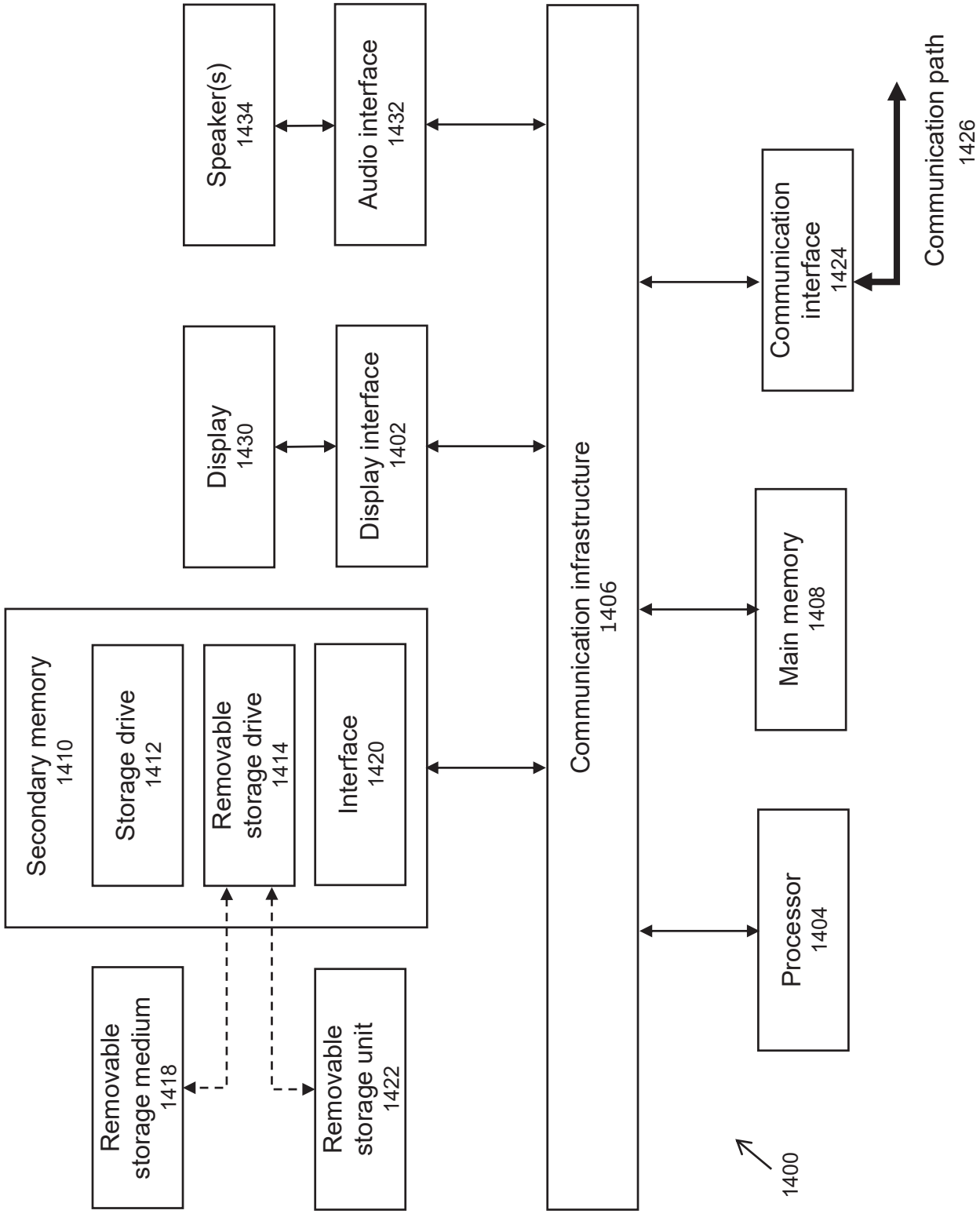

Speaker(s)
1434
Audio interface
1432
Display
1430
Display interface
1402
Secondary memory
1410
Storage drive
1412
Removable storage drive
1414
Interface
1420
Communication infrastructure
1406
Main memory
1408
Communication interface
1424
Communication path
1426
Removable storage medium
1418
Removable storage unit
1422
Processor
1404
1400

Figure 14