



(19)中華民國智慧財產局

(12)發明說明書公告本

(11)證書號數：TW I714834 B

(45)公告日：中華民國 110 (2021) 年 01 月 01 日

(21)申請案號：107107233

(22)申請日：中華民國 107 (2018) 年 03 月 05 日

(51)Int. Cl. : G06N3/08 (2006.01)

G06N20/00 (2019.01)

G06T7/00 (2017.01)

(30)優先權：2017/06/07 中國大陸

201710421333.5

(71)申請人：開曼群島商創新先進技術有限公司(開曼群島) ADVANCED NEW TECHNOLOGIES CO., LTD. (KY)

開曼群島

(72)發明人：馬晨光 (CN)

(74)代理人：林志剛

(56)參考文獻：

TW 201403547A

TW 201629851A

CN 106599772A

CN 106778525A

Jianwei Yang 等 3 人，" Learn Convolutional Neural Network for Face Anti-Spoofing"，2014/8/26，https://www.researchgate.net/publication/265089659_Learn_Convolutional_Neural_Network_for_Face_Anti-Spoofing/download

265089659_Learn_Convolutional_Neural_Network_for_Face_Anti-Spoofing/download

David Menotti 等 7 人，" Deep Representations for Iris, Face, and Fingerprint Spoofing Attack Detection"，2014/10，https://www.researchgate.net/publication/266676858_Deep_Representations_for_Iris_Face_and_Fingerprint_Spoofing_Attack_Detection/download

266676858_Deep_Representations_for_Iris_Face_and_Fingerprint_Spoofing_Attack_Detection/download

審查人員：許人偉

申請專利範圍項數：15 項 圖式數：5 共 37 頁

(54)名稱

人臉活體檢測方法、裝置以及電子設備

(57)摘要

本發明揭示了一種人臉活體檢測方法、裝置以及電子設備。所述方法包括：根據人臉採集全域影像，訓練第一深度學習模型；以及，根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型；根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。利用本發明實施例，有利於提取更多的人臉活體影像特徵，訓練後的所述第一深度學習模型和所述第二深度學習模型聯合而言，相比於現有技術的模型對於人臉活體的特徵表達能力更好，進而有利於提高人臉活體檢測結果的準確性。

指定代表圖：

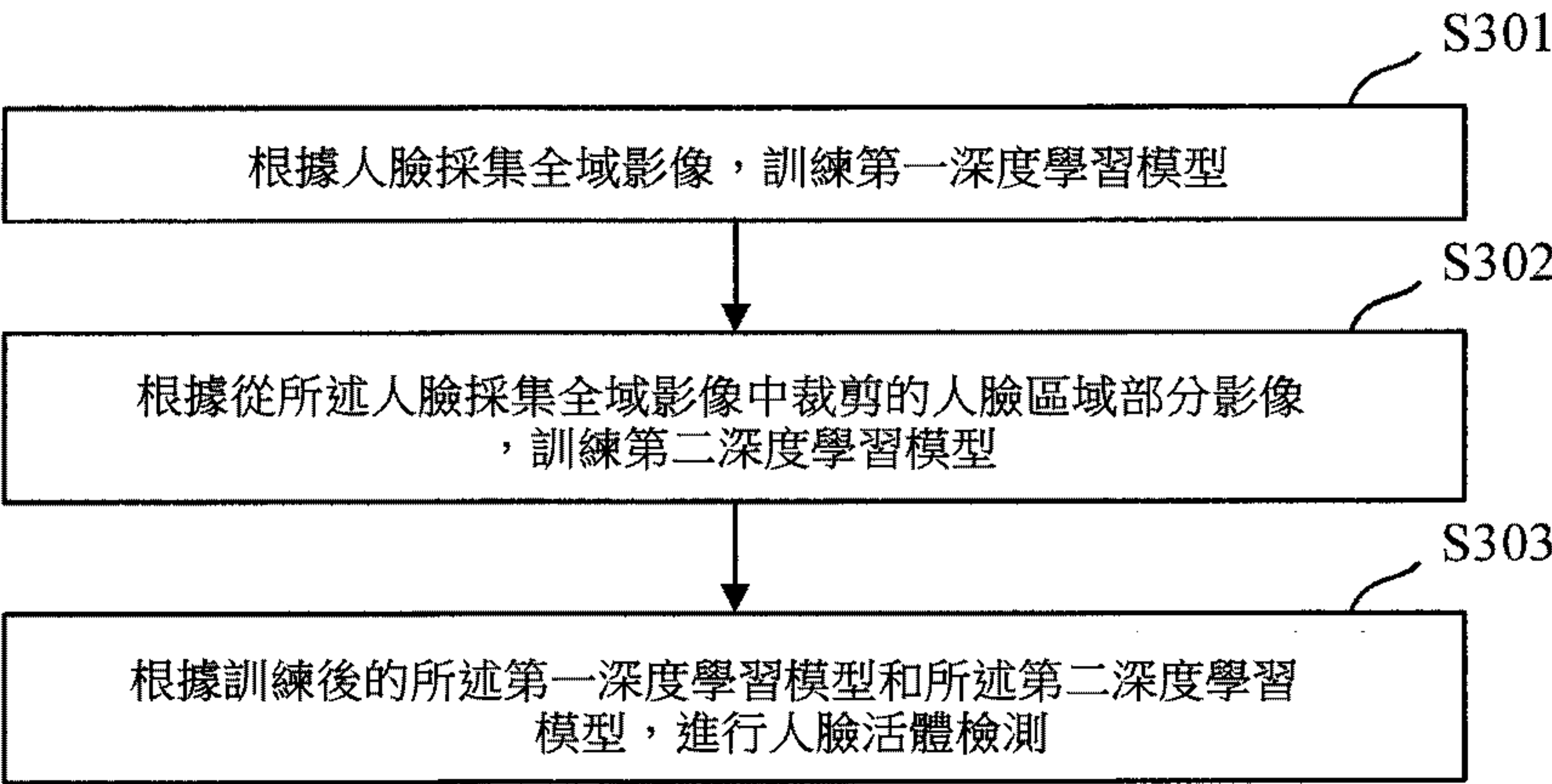


圖 3



I714834

【發明摘要】

【中文發明名稱】

人臉活體檢測方法、裝置以及電子設備

【中文】

本發明揭示了一種人臉活體檢測方法、裝置以及電子設備。所述方法包括：根據人臉採集全域影像，訓練第一深度學習模型；以及，根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型；根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。利用本發明實施例，有利於提取更多的人臉活體影像特徵，訓練後的所述第一深度學習模型和所述第二深度學習模型聯合而言，相比於現有技術的模型對於人臉活體的特徵表達能力更好，進而有利於提高人臉活體檢測結果的準確性。

【指定代表圖】第(3)圖。
【代表圖之符號簡單說明】無

【特徵化學式】無

【發明說明書】

【中文發明名稱】

人臉活體檢測方法、裝置以及電子設備

【技術領域】

本發明關於電腦軟體技術領域，尤其關於一種人臉活體檢測方法、裝置以及電子設備。

【先前技術】

人臉活體檢測技術用於判斷目前人臉識別的操作者是否為活體用戶，進而可以實現對屏幕重放攻擊、打印照片攻擊、三維建模攻擊等非活體攻擊進行攔截。

目前，人臉活體檢測技術可分為兩大類：第一類技術為主動式人臉活體檢測技術，這類技術要求用戶在進行人臉識別時根據指令，配合完成特定的活體動作，如眨眼、轉頭、張嘴等，從而活體檢測模組可以根據操作者是否準確地完成了活體動作，判定操作者是否是活體；第二類技術為被動式人臉活體檢測技術，這類技術不需要用戶配合完成活體動作，用戶體驗較好，但技術難度高，主要依賴輸入的單幀影像資訊或者其他設備傳感器資訊進行活體檢測。

在現有技術中，對於上述的被動式人臉活體檢測技術，通常使用活體與非活體的人臉採集影像對於單一的深度學習模型進行監督式訓練，然後使用訓練後的模型針對

輸入的單幀影像進行人臉活體預測。

但是，這種技術方案嚴重依賴訓練資料的人臉非活體攻擊種類，同時受限於訓練資料不夠豐富的客觀條件，難以充分提取人臉活體影像特徵，從而使得其模型對於人臉活體的特徵表達能力不足，進而會降低人臉活體檢測結果的準確性。

【發明內容】

本發明實施例提供一種人臉活體檢測方法、裝置以及電子設備，用以解決現有技術中的如下技術問題：基於單一的深度學習模型的技術方案難以充分提取人臉活體影像特徵，從而使得其模型對於人臉活體的特徵表達能力不足，進而會降低人臉活體檢測結果的準確性。

為解決上述技術問題，本發明實施例是這樣實現的：

本發明實施例提供的一種人臉活體檢測方法，包括：

根據人臉採集全域影像，訓練第一深度學習模型；以及，

根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型；

根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。

本發明實施例提供的一種人臉活體檢測裝置，包括：

訓練模組，根據人臉採集全域影像，訓練第一深度學習模型；以及，

根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型；

檢測模組，根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。

本發明實施例採用的上述至少一個技術方案能夠達到以下有益效果：有利於提取更多的人臉活體影像特徵，訓練後的所述第一深度學習模型和所述第二深度學習模型聯合而言，相比於現有技術的模型對於人臉活體的特徵表達能力更好，進而有利於提高人臉活體檢測結果的準確性，因此，可以部分或全部地解決現有技術中的問題。

【圖式簡單說明】

為了更清楚地說明本發明實施例或現有技術中的技術方案，下面將對實施例或現有技術描述中所需要使用的附圖作簡單地介紹，顯而易見地，下面描述中的附圖僅僅是本發明中記載的一些實施例，對於本領域普通技術人員來講，在不付出創造性勞動性的前提下，還可以根據這些附圖獲得其他的附圖。

圖1為本發明的方案中包含的模型訓練階段的示例性流程示意圖；

圖2為本發明的方案中包含活體檢測階段的示例性流程示意圖；

圖3為本發明實施例提供的一種人臉活體檢測方法的流程示意圖；

圖4為本發明實施例提供一種人臉採集全域影像和人臉區域部分影像的對比示意圖；

圖5為本發明實施例提供的對應於圖3的一種人臉活體檢測裝置的結構示意圖。

【實施方式】

本發明實施例提供一種人臉活體檢測方法、裝置以及電子設備。

為了使本技術領域的人員更好地理解本發明中的技術方案，下面將結合本發明實施例中的附圖，對本發明實施例中的技術方案進行清楚、完整地描述，顯然，所描述的實施例僅僅是本發明一部分實施例，而不是全部的實施例。基於本發明中的實施例，本領域普通技術人員在沒有作出創造性勞動前提下所獲得的所有其他實施例，都應當屬於本發明保護的範圍。

本發明的方案中的深度學習模型均是基於神經網路的，為了便於描述，先基於一個示例，結合圖1和圖2對本發明的方案的核心思想進行闡述：

在該示例下，本發明的方案可以分為模型訓練階段與活體檢測階段；

圖1為本發明的方案中包含的模型訓練階段的示例性流程示意圖，在模型訓練階段，使用人臉採集影像中的活體與非活體樣本(屬於訓練資料集)訓練兩種獨立的深度學習模型：第一深度學習模型、第二深度學習模型；其中，

第一深度學習模型的輸入影像是採集的人臉採集全域影像，第二深度學習模型的輸入影像可以是在該人臉採集全域影像中裁剪的人臉區域部分影像，第一深度學習模型與第二深度學習模型可以採用不同的深度學習網路結構(亦即，模型所基於的神經網路的結構)，不同的網路結構對於不同的影像特徵可以具有不同的敏感性；使用活體與非活體訓練資料集，基於深度學習方法，對第一深度學習模型和第二深度學習模型訓練完畢；

圖2為本發明的方案中包含活體檢測階段的示例性流程示意圖，在活體檢測階段，採集用戶刷臉影像作為該用戶的人臉採集全域影像，並輸入第一深度學習模型，得到預測得分 PA ；同時，在該用戶刷臉影像上執行人臉檢測，根據檢測結果從該用戶刷臉影像中裁剪出人臉區域部分影像輸入第二深度學習模型，得到預測得分 PB ；然後，比如可以透過將 $PA+PB$ 的預測得分結果同設定好的閾值(比如，閾值可以為1)進行對比，進行聯合決策確定針對該用戶刷臉影像的人臉活體檢測結果。

基於上面的核心思想，下面對本發明的方案進行詳細說明。

圖3為本發明實施例提供的一種人臉活體檢測方法的流程示意圖。從程式角度而言，該流程的執行主體可以是伺服器上或者終端上的程式，比如，身份認證程式、電子商務應用等。從設備角度而言，該流程的執行主體包括但不限於可作為伺服器或者終端的以下至少一種設備：門禁

設備、個人電腦、中型電腦、電腦集群、手機、平板電腦、智慧型穿戴式設備、車機、銷售終端(Point Of Sale，POS)機等。

圖3中的流程可以包括以下步驟：

S301：根據人臉採集全域影像，訓練第一深度學習模型。

在本發明實施例中，訓練第一深度學習模型所使用的人臉採集全域影像可以有多個樣本，在這多個樣本中，有的樣本是活體採集影像，透過對人臉活體進行拍攝而採集得到，可以作為正樣本，有的樣本是非活體採集影像，透過對諸如人臉圖片、人臉模型等非活體進行拍攝而採集得到，可以作為負樣本。

在本發明實施例中，第一深度學習模型是分類模型，人臉採集全域影像作為該分類模型的輸入，經過模型處理後，至少可以將人臉採集全域影像分類至活體採集影像類別或者非活體採集影像類別。訓練第一深度學習模型的目的在於：提高第一深度學習模型的分類準確性。

S302：根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型。

在本發明實施例中，一般地，人臉採集全域影像除了包含整個人臉區域以外，往往還包含一些無關區域，無關區域比如可以是背景區域、人體除人臉以外的區域等。而人臉區域部分影像可以不包含無關區域，而是只包含至少部分人臉區域，比如，整個人臉區域，或者眼部區域，或

者鼻部區域等。第二深度學習模型可以有一個或者多個，每個第二深度學習模型可以分別對應於一類人臉區域。

圖4為本發明實施例提供一種人臉採集全域影像和人臉區域部分影像的對比示意圖。

在圖4中，(a)為人臉採集全域影像，為了便於理解，在(a)中用虛線對人臉區域部分進行了標記，對(a)相應地進行裁剪可以得到如(b)所示的人臉區域部分影像。

另外，在人臉區域部分影像為只包含部分人臉區域的影像的情況下，人臉採集全域影像也可以是包含整個人臉區域且基本不包含無關區域的影像。

在本發明實施例中，訓練第二深度學習模型所使用的人臉區域部分影像也可以有多個樣本，在這多個樣本中，有的樣本是活體採集影像，可以作為正樣本，有的樣本是非活體採集影像，可以作為負樣本。

在本發明實施例中，第二深度學習模型也是分類模型，人臉區域部分影像作為該分類模型的輸入，經過模型處理後，至少可以將人臉區域部分影像分類至活體採集影像類別或者非活體採集影像類別。訓練第二深度學習模型的目的在於：提高第二深度學習模型的分類準確性。

人臉區域部分影像除了可以從人臉採集全域影像裁剪得到以外，也可以不依賴於人臉採集全域影像而專門採集得到的。

在本發明實施例中，第一深度學習模型與第二深度學習模型在被訓練前可以是不同的模型，也可以是相同的模

型。

本發明對步驟 S301 和 S302 的執行順序並不做限定，可以同時並行執行，也可以先後執行等。

S303：根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。

圖 3 中的各步驟的執行主體可以是同一設備或同一程式，也可以是不同設備或不同程式。比如，步驟 S301~S303 的執行主體均為設備 1；又比如，步驟 S301~S302 的執行主體為設備 1，步驟 S303 的執行主體為設備 2；等等。

透過圖 3 的方法，有利於提取更多的人臉活體影像特徵，訓練後的所述第一深度學習模型和所述第二深度學習模型聯合而言，相比於現有技術的模型對於人臉活體的特徵表達能力更好，進而有利於提高人臉活體檢測結果的準確性，因此，可以部分或全部地解決現有技術中的問題。

基於圖 3 的方法，本發明實施例還提供了該方法的一些具體實施方案，以及擴展方案，下面進行說明。

在本發明實施例中，為了便於實現第一深度學習模型與第二深度學習模型對影像特徵的敏感性差異，第一深度學習模型與第二深度學習模型優選地可以採用不同的深度學習網路結構。

兩種深度學習模型的網路結構不同具體可以指：這兩種深度學習模型的一個或多個網路結構參數不同，其中，

網路結構參數比如可以包括：隱變量層數、隱變量層類型、神經元節點數、輸入層節點數、輸出層節點數等。

當然，對於一些特定的深度學習模型也可能有相應的特定參數。比如，對於目前在影像領域應用廣泛的基於卷積神經網路的深度學習模型，卷積單元的卷積核大小也屬於這種深度學習模型特定的網路結構參數。

對於本發明的方案，一般地，所述不同的深度學習網路結構的以下至少一種參數不同：隱變量層數、隱變量層類型、神經元節點數、卷積單元的卷積核大小等。

在本發明實施例中，為了提高模型訓練效率和模型訓練可靠性，可以採用有監督式的訓練方式進行模型訓練。

例如，在監督式的訓練方式下，對於步驟S301，所述人臉採集全域影像具有第一標籤，所述第一標籤表明其對應的人臉採集全域影像是活體採集影像還是非活體採集影像；

所述根據人臉採集全域影像，訓練第一深度學習模型，具體可以包括：將人臉採集全域影像輸入第一深度學習模型，所述第一深度學習模型提取所述人臉採集全域影像的特徵，並根據提取的特徵，預測所述人臉採集全域影像是活體採集影像還是非活體採集影像；根據預測結果和所述人臉採集全域影像的所述第一標籤，對所述第一深度學習模型進行調整。一般地，當預測結果與第一標籤不符時，對第一深度學習模型進行調整，以使調整後的第一深度學習模型能夠重新預測得到與第一標籤相符的預測結

果。

其中，第一深度學習模型在訓練過程中提取的特徵較佳地可以包括人臉採集全域影像的影像結構性特徵，比如，人臉採集全域影像中包含的屏幕照片邊緣、人臉畸變等特徵。

又例如，類似地，在監督式的訓練方式下，對於步驟S302，所述人臉區域部分影像具有第二標籤，所述第二標籤表明其對應的人臉區域部分影像是活體採集影像還是非活體採集影像；

所述根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型，具體可以包括：獲取從所述人臉採集全域影像中裁剪的人臉區域部分影像；將所述獲取的人臉區域部分影像輸入第二深度學習模型，所述第二深度學習模型提取所述人臉區域部分影像的特徵，並根據提取的特徵，預測所述人臉區域部分影像是活體採集影像還是非活體採集影像；根據預測結果和所述人臉區域部分影像的所述第二標籤，對所述第二深度學習模型進行調整。一般地，當預測結果與第二標籤不符時，對第二深度學習模型進行調整，以使調整後的第二深度學習模型能夠重新預測得到與第二標籤相符的預測結果。

其中，第二深度學習模型在訓練過程中提取的特徵較佳地可以包括人臉區域部分影像的影像材質性特徵，比如，人臉區域部分影像中包含的模糊、紋理、顏色失真等特徵。

在上面的兩個例子中，第一深度學習模型與第二深度學習模型對不同的影像特徵具有不同的敏感性，第一深度學習模型對於影像結構性特徵更敏感，第二深度學習模型對於影像材質性特徵更敏感。對於人臉影像而言，影像結構性特徵是相對全域和概括的特徵，影像材質性特徵是相對局部和細化的特徵。

如此，聯合訓練後的第一深度學習模型和第二深度學習模型，可以對人臉影像進行更有層次和更豐富的特徵提取，進而也可以有利於聯合決策得到更準確的人臉活體檢測結果。

在本發明實施例中，透過使對應的訓練資料集不同和/或對應的深度學習網路結構不同，可能實現第一深度學習模型與第二深度學習模型對不同的影像特徵的不同的敏感性。

例如，若第一深度學習模型與第二深度學習模型基於卷積神經網路；第一深度學習模型所基於的卷積神經網路的卷積單元的卷積核可以較大，以便於第一深度學習模型提取人臉採集全域影像的影像結構性特徵；相對地，第二深度學習模型所基於的卷積神經網路的卷積單元的卷積核可以較小，以便於第二深度學習模型提取人臉區域部分影像的影像材質性特徵。因此，在該例的情況下，第一深度學習模型所基於的卷積神經網路的卷積單元的卷積核大於第二深度學習模型所基於的卷積神經網路的卷積單元的卷積核。

需要說明的是，卷積核大小僅是一種可影響所述敏感性的參數示例，其他的網路結構參數也可能影響所述敏感性。

在本發明實施例中，對於步驟S303，可以透過聯合訓練後的第一深度學習模型和第二深度學習模型進行決策，進行人臉活體檢測。決策的具體方式可以有多種：比如，分別用第一深度學習模型和第二深度學習模型進行單獨決策，再綜合各單獨決策結果以確定最終決策結果；再比如，可以先用第一深度學習模型和第二深度學習模型其中之一進行單獨決策，當單獨決策結果符合一定條件時，可以直接作為最終決策結果，否則，再結合剩下的另一模型綜合地進行決策，得到最終決策結果；等等。

若採用上一段中列舉的第一種方式，舉例如下：

例如，對於步驟S303，所述根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測，具體可以包括：獲取為進行人臉活體檢測採集的人臉採集全域影像（通常即為：用戶刷臉影像）；將所述採集的人臉採集全域影像輸入訓練後的所述第一深度學習模型進行處理，得到對應的第一預測資料；以及，獲取從所述採集的人臉採集全域影像中裁剪的人臉區域部分影像，並輸入訓練後的所述第二深度學習模型進行處理，得到對應的第二預測資料；透過根據所述第一預測資料和所述第二預測資料進行聯合決策，得到針對所述用戶刷臉影像的人臉活體檢測結果。

第一預測資料比如可以是上述的預測得分 **PA**，第二預測資料比如可以是上述的預測得分 **PB**。當然，預測得分只是第一預測資料、第二預測資料的一種表現形式示例，還可以有其他的表現形式，比如，機率值、布林型值等。

上面為本發明實施例提供的一種人臉活體檢測方法，基於同樣的發明思路，本發明實施例還提供了對應的裝置，如圖5所示。

圖5為本發明實施例提供的對應於圖3的一種人臉活體檢測裝置的結構示意圖，該裝置可以位於圖3中流程的執行主體上，包括：

訓練模組501，根據人臉採集全域影像，訓練第一深度學習模型；以及，

根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型；

檢測模組502，根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。

可選地，所述第一深度學習模型與所述第二深度學習模型採用不同的深度學習網路結構。

可選地，所述不同的深度學習網路結構的以下至少一種參數不同：隱變量層數、隱變量層類型、神經元節點數、卷積單元的卷積核大小。

可選地，所述人臉採集全域影像具有第一標籤，所述第一標籤表明其對應的人臉採集全域影像是活體採集影像還是非活體採集影像；

所述訓練模組501根據人臉採集全域影像，訓練第一深度學習模型，具體包括：

所述訓練模組501將人臉採集全域影像輸入第一深度學習模型，該第一深度學習模型根據所述人臉採集全域影像的影像結構性特徵，預測所述人臉採集全域影像是活體採集影像還是非活體採集影像；

根據預測結果和所述人臉採集全域影像的所述第一標籤，對所述第一深度學習模型進行調整。

可選地，所述人臉區域部分影像具有第二標籤，所述第二標籤表明其對應的人臉區域部分影像是活體採集影像還是非活體採集影像；

所述訓練模組501根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型，具體包括：

所述訓練模組501獲取從所述人臉採集全域影像中裁剪的人臉區域部分影像；將所述獲取的人臉區域部分影像輸入第二深度學習模型，該第二深度學習模型根據所述人臉區域部分影像的影像材質性特徵，預測所述人臉區域部分影像是活體採集影像還是非活體採集影像；

根據預測結果和所述人臉區域部分影像的所述第二標籤，對所述第二深度學習模型進行調整。

可選地，所述第一深度學習模型與所述第二深度學習模型基於卷積神經網路；

所述第一深度學習模型所基於的卷積神經網路的卷積

單元的卷積核大於所述第二深度學習模型所基於的卷積神經網路的卷積單元的卷積核，以便於所述第一深度學習模型提取所述人臉採集全域影像的影像結構性特徵，以及以便於所述第二深度學習模型提取所述人臉區域部分影像的影像材質性特徵。

可選地，所述檢測模組502根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測，具體包括：

所述檢測模組502獲取為進行人臉活體檢測採集的人臉採集全域影像；

將所述採集的人臉採集全域影像輸入訓練後的所述第一深度學習模型進行處理，得到對應的第一預測資料；以及，

獲取從所述採集的人臉採集全域影像中裁剪到人臉區域部分影像，並輸入訓練後的所述第二深度學習模型進行處理，得到對應的第二預測資料；

透過根據所述第一預測資料和所述第二預測資料進行聯合決策，確定針對所述用戶刷臉影像的人臉活體檢測結果。

基於同樣的發明思路，本發明實施例還提供了對應的一種電子設備，包括：

至少一個處理器；以及，

與所述至少一個處理器通訊連接的記憶體；其中，

所述記憶體儲存有可被所述至少一個處理器執行的指

令，所述指令被所述至少一個處理器執行，以使所述至少一個處理器能夠：

根據人臉採集全域影像，訓練第一深度學習模型；以及，

根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型；

根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。

基於同樣的發明思路，本發明實施例還提供了對應的一種非易失性電腦儲存媒體，儲存有電腦可執行指令，所述電腦可執行指令設置為：

根據人臉採集全域影像，訓練第一深度學習模型；以及，

根據從所述人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型；

根據訓練後的所述第一深度學習模型和所述第二深度學習模型，進行人臉活體檢測。

本說明書中的各個實施例均採用遞進的方式來描述，各個實施例之間相同相似的部分互相參見即可，每個實施例重點說明的都是與其他實施例的不同之處。尤其，對於裝置、電子設備、非易失性電腦儲存媒體實施例而言，由於其基本相似於方法實施例，所以描述的比較簡單，相關之處參見方法實施例的部分說明即可。

本發明實施例提供的裝置、電子設備、非易失性電腦

儲存媒體與方法是對應的，因此，裝置、電子設備、非易失性電腦儲存媒體也具有與對應方法類似的有益技術效果，由於上面已經對方法的有益技術效果進行了詳細說明，因此，這裡不再贅述對應裝置、電子設備、非易失性電腦儲存媒體的有益技術效果。

在20世紀90年代，對於一個技術的改進可以很明顯地區分是硬體上的改進(例如，對二極體、電晶體、開關等電路結構的改進)還是軟體上的改進(對於方法流程的改進)。然而，隨著技術的發展，當今的很多方法流程的改進已經可以視為硬體電路結構的直接改進。設計人員幾乎都透過將改進的方法流程編程到硬體電路中來得到相應的硬體電路結構。因此，不能說一個方法流程的改進就不能用硬體實體模組來實現。例如，可編程邏輯裝置(Programmable Logic Device, PLD)(例如現場可編程閘陣列(Field Programmable Gate Array, FPGA))就是這樣一種積體電路，其邏輯功能由用戶對裝置編程來確定。由設計人員自行編程來把一個數位系統“整合”在一片PLD上，而不需要請晶片製造廠商來設計和製作專用的積體電路晶片。而且，如今，取代手工地製作積體電路晶片，這種編程也多半改用“邏輯編譯器(logic compiler)”軟體來實現，它與程式開發撰寫時所用的軟體編譯器相類似，而要編譯之前的原始代碼也得用特定的程式語言來撰寫，此稱之為硬體描述語言(Hardware Description Language, HDL)，而HDL也並非僅有一種，而是有許多種，如ABEL(Advanced

Boolean Expression Language)、AHDL(Altera Hardware Description Language)、Confluence、CUPL(Cornell University Programming Language)、HDCal、JHDL(Java Hardware Description Language)、Lava、Lola、MyHDL、PALASM、RHDL(Ruby Hardware Description Language)等，目前最普遍使用的是 VHDL(Very-High-Speed Integrated Circuit Hardware Description Language) 與 Verilog。本領域技術人員也應該清楚，只需要將方法流程用上述幾種硬體描述語言稍作邏輯編程並編程到積體電路中，就可以很容易得到實現該邏輯方法流程的硬體電路。

控制器可以按任何適當的方式實現，例如，控制器可以採取例如微處理器或處理器以及儲存可由該(微)處理器執行的電腦可讀程式代碼(例如軟體或韌體)的電腦可讀媒體、邏輯閘、開關、專用積體電路(Application Specific Integrated Circuit, ASIC)、可編程邏輯控制器和嵌入式微控制器的形式，控制器的例子包括但不限於以下微控制器：ARC 625D、Atmel AT91SAM、Microchip PIC18F26K20 以及Silicone Labs C8051F320，記憶體控制器還可以被實現為記憶體的控制邏輯的一部分。本領域技術人員也知道，除了以純電腦可讀程式代碼方式實現控制器以外，完全可以透過將方法步驟進行邏輯編程來使得控制器以邏輯閘、開關、專用積體電路、可編程邏輯控制器和嵌入式微控制器等的形式來實現相同功能。因此這種控制器可以被認為是一種硬體部件，而對其內包括的用於實

現各種功能的裝置也可以視為硬體部件內的結構。或者甚至，可以將用於實現各種功能的裝置視為既可以是實現方法的軟體模組又可以是硬體部件內的結構。

上述實施例闡明的系統、裝置、模組或單元，具體可以由電腦晶片或實體來實現，或者由具有某種功能的產品來實現。一種典型的實現設備為電腦。具體地，電腦例如可以為個人電腦、膝上型電腦、蜂窩電話、相機電話、智慧型電話、個人數位助理、媒體播放器、導航設備、電子郵件設備、遊戲控制台、平板電腦、穿戴式設備或者這些設備中的任何設備的組合。

為了描述的方便，描述以上裝置時以功能分為各種單元分別描述。當然，在實施本發明時可以把各單元的功能在同一個或多個軟體和/或硬體中實現。

本領域內的技術人員應明白，本發明的實施例可提供為方法、系統、或電腦程式產品。因此，本發明可採用完全硬體實施例、完全軟體實施例、或結合軟體和硬體方面的實施例的形式。而且，本發明可採用在一個或多個其中包含有電腦可用程式代碼的電腦可用儲存媒體(包括但不限於磁碟記憶體、**CD-ROM**、光學記憶體等)上實施的電腦程式產品的形式。

本發明是參照根據本發明實施例的方法、設備(系統)、和電腦程式產品的流程圖和/或方塊圖來描述的。應理解可由電腦程式指令實現流程圖和/或方塊圖中的每一流程和/或方塊、以及流程圖和/或方塊圖中的流程和/或方

塊的結合。可提供這些電腦程式指令到通用電腦、專用電腦、嵌入式處理機或其他可編程資料處理設備的處理器以產生一個機器，使得透過電腦或其他可編程資料處理設備的處理器執行的指令產生用於實現在流程圖一個流程或多個流程和/或方塊圖一個方塊或多個方塊中指定的功能的裝置。

這些電腦程式指令也可被儲存在能引導電腦或其他可編程資料處理設備以特定方式操作的電腦可讀記憶體中，使得儲存在該電腦可讀記憶體中的指令產生包括指令裝置的製造品，該指令裝置實現在流程圖一個流程或多個流程和/或方塊圖一個方塊或多個方塊中指定的功能。

這些電腦程式指令也可被裝載到電腦或其他可編程資料處理設備上，使得在電腦或其他可編程設備上執行一系列操作步驟以產生電腦實現的處理，從而在電腦或其他可編程設備上執行的指令提供用於實現在流程圖一個流程或多個流程和/或方塊圖一個方塊或多個方塊中指定的功能的步驟。

在一個典型的配置中，計算設備包括一個或多個處理器(CPU)、輸入/輸出介面、網路介面和內部記憶體。

內部記憶體可能包括電腦可讀媒體中的非永久性記憶體，隨機存取記憶體(RAM)和/或非易失性內部記憶體等形式，如唯讀記憶體(ROM)或閃存(flash RAM)。內部記憶體是電腦可讀媒體的示例。

電腦可讀媒體包括永久性和非永久性、可移動和非可

移動媒體可以由任何方法或技術來實現資訊儲存。資訊可以是電腦可讀指令、資料結構、程式的模組或其他資料。電腦的儲存媒體的例子包括，但不限於相變內部記憶體 (PRAM)、靜態隨機存取記憶體 (SRAM)、動態隨機存取記憶體 (DRAM)、其他類型的隨機存取記憶體 (RAM)、唯讀記憶體 (ROM)、電可擦除可編程唯讀記憶體 (EEPROM)、快閃記憶體或其他內部記憶體技術、唯讀光碟唯讀記憶體 (CD-ROM)、數位多功能光碟 (DVD) 或其他光學儲存、磁盒式磁帶，磁帶磁碟儲存或其他磁性儲存設備或任何其他非傳輸媒體，可用於儲存可以被計算設備訪問的資訊。按照本文中的界定，電腦可讀媒體不包括暫態性電腦可讀媒體 (transitory media)，如調變的資料信號和載波。

還需要說明的是，術語“包括”、“包含”或者其任何其他變體意在涵蓋非排他性的包含，從而使得包括一系列要素的過程、方法、商品或者設備不僅包括那些要素，而且還包括沒有明確列出的其他要素，或者是還包括為這種過程、方法、商品或者設備所固有的要素。在沒有更多限制的情況下，由語句“包括一個……”限定的要素，並不排除在包括所述要素的過程、方法、商品或者設備中還存在另外的相同要素。

本發明可以在由電腦執行的電腦可執行指令的一般上下文中描述，例如程式模組。一般地，程式模組包括執行特定任務或實現特定抽象資料類型的例程、程式、對象、組件、資料結構等等。也可以在分散式計算環境中實踐本

發明，在這些分散式計算環境中，由透過通訊網路而被連接的遠程處理設備來執行任務。在分散式計算環境中，程式模組可以位於包括儲存設備在內的本地和遠程電腦儲存媒體中。

本說明書中的各個實施例均採用遞進的方式描述，各個實施例之間相同相似的部分互相參見即可，每個實施例重點說明的都是與其他實施例的不同之處。尤其，對於系統實施例而言，由於其基本相似於方法實施例，所以描述的比較簡單，相關之處參見方法實施例的部分說明即可。

以上所述僅為本發明的實施例而已，並不用於限制本發明。對於本領域技術人員來說，本發明可以有各種更改和變化。凡在本發明的精神和原理之內所作的任何修改、等同替換、改進等，均應包含在本發明的申請專利範圍之內。

【符號說明】

S301、S302、S303：方法步驟

501：訓練模組

502：檢測模組

【發明申請專利範圍】

【第 1 項】

一種人臉活體檢測方法，包括：

根據人臉採集全域影像，藉由對多個全域影像進行監督式的訓練來訓練第一深度學習模型，其中，該全域影像包括藉由拍攝活體人臉所採集並標記為正樣本的活體人臉影像，以及藉由拍攝人臉圖片或人臉模型的非活體人臉影像所採集並標記為負樣本的非活體人臉影像，該人臉採集全域影像包括人臉區域和背景區域，該第一深度學習模型是分類模型，提取該人臉採集全域影像的結構性特徵，並將人臉採集全域影像分類至活體採集影像類別或者非活體採集影像類別，其中，該結構性特徵包括至少以下任一種：屏幕照片邊緣以及人臉畸變；

根據從該人臉採集全域影像中裁剪的人臉區域部分影像，藉由對多個裁剪的人臉區域部分影像進行監督式的訓練來訓練第二深度學習模型，其中，該第二深度學習模型包括分別與該人臉區域的鼻部對應的鼻部深度學習模型和與眼部對應的眼部深度學習模型，該人臉區域部分影像包括人臉區域但不包括背景區域，該第二深度學習模型是分類模型，提取該人臉區域部分影像的影像材質性特徵，並將該人臉區域部分影像分類至活體採集影像類別或者非活體採集影像類別，其中，該影像材質性特徵包括至少以下任一種：模糊、紋理及顏色失真；以及

根據訓練後的該第一深度學習模型對第一全域影像進

行人臉活體檢測以獲得第一預測得分，以及使用多個訓練後的該第二深度學習模型獲取第二預測得分，包括：

獲取採集的第一全域影像用於人臉活體檢測；

將該第一全域影像輸入訓練後的該第一深度學習模型中進行處理以獲得該第一預測得分；

獲取從該第一全域影像裁剪的多個提取的人臉影像，該提取的人臉影像包括眼部影像區域影像和鼻部影像區域影像，並將該提取的人臉影像輸入到訓練後的包括該眼部深度學習模型和該鼻部深度學習模型的該第二深度學習模型中進行處理，以獲得第二預測得分；

基於該第一預測得分和該第二預測得分，產生預測得分結果；以及

比較該預測得分結果與閾值，以確定該第一全域影像是活體影像還是非活體影像。

【第2項】

如申請專利範圍第1項所述的方法，該第一深度學習模型與該第二深度學習模型採用不同的深度學習網路結構。

【第3項】

如申請專利範圍第2項所述的方法，該不同的深度學習網路結構的以下至少一種參數不同：隱變量層數、隱變量層類型、神經元節點數、卷積單元的卷積核大小。

【第4項】

如申請專利範圍第1項所述的方法，該人臉採集全域影像具有第一標籤，該第一標籤表明其對應的人臉採集全域影像是活體採集影像還是非活體採集影像；

該根據人臉採集全域影像，訓練第一深度學習模型，具體包括：

將人臉採集全域影像輸入第一深度學習模型，該第一深度學習模型根據該人臉採集全域影像的影像結構性特徵，預測該人臉採集全域影像是活體採集影像還是非活體採集影像；以及

根據預測結果和該人臉採集全域影像的該第一標籤，對該第一深度學習模型進行調整。

【第5項】

如申請專利範圍第1項所述的方法，該人臉區域部分影像具有第二標籤，該第二標籤表明其對應的人臉區域部分影像是活體採集影像還是非活體採集影像；

該根據從該人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型，具體包括：

獲取從該人臉採集全域影像中裁剪的人臉區域部分影像；

將該獲取的人臉區域部分影像輸入第二深度學習模型，該第二深度學習模型根據該人臉區域部分影像的影像材質性特徵，預測該人臉區域部分影像是活體採集影像還是非活體採集影像；以及

根據預測結果和該人臉區域部分影像的該第二標籤，

對該第二深度學習模型進行調整。

【第6項】

如申請專利範圍第1項所述的方法，該第一深度學習模型與該第二深度學習模型基於卷積神經網路；以及

該第一深度學習模型所基於的卷積神經網路的卷積單元的卷積核大於該第二深度學習模型所基於的卷積神經網路的卷積單元的卷積核。

【第7項】

如申請專利範圍第1項所述的方法，該根據訓練後的該第一深度學習模型和該第二深度學習模型，進行人臉活體檢測，具體包括：

獲取為進行人臉活體檢測採集的人臉採集全域影像；

將該採集的人臉採集全域影像輸入訓練後的該第一深度學習模型進行處理，得到對應的第一預測資料；

獲取從該採集的人臉採集全域影像中裁剪的人臉區域部分影像，並輸入訓練後的該第二深度學習模型進行處理，得到對應的第二預測資料；以及

透過根據該第一預測資料和該第二預測資料進行聯合決策，得到針對該用戶刷臉影像的人臉活體檢測結果。

【第8項】

一種人臉活體檢測裝置，包括：

訓練模組，根據人臉採集全域影像，藉由對多個全域影像進行監督式的訓練訓練第一深度學習模型，以及，根據從該人臉採集全域影像中裁剪的人臉區域部分影像，藉

由對多個裁剪的人臉區域部分影像進行監督式的訓練來訓練第二深度學習模型；以及

檢測模組，根據訓練後的該第一深度學習模型對第一全域影像進行人臉活體檢測，以獲得第一預測得分，以及使用多個訓練後的該第二深度學習模型獲取第二預測得分，包括：

獲取採集的第一全域影像用於人臉活體檢測；

將該第一全域影像輸入訓練後的該第一深度學習模型中進行處理以獲得該第一預測得分；

獲取從該第一全域影像裁剪的多個提取的人臉影像，該提取的人臉影像包括眼部影像區域影像和鼻部影像區域影像，並將該提取的人臉影像輸入到訓練後的包括該眼部深度學習模型和該鼻部深度學習模型的該第二深度學習模型中進行處理，以獲得第二預測得分；

基於該第一預測得分和該第二預測得分，產生預測得分結果；以及

比較該預測得分結果與閾值，以確定該第一全域影像是活體影像還是非活體影像，

其中，該全域影像包括藉由拍攝活體人臉所採集並標記為正樣本的活體人臉影像，以及藉由拍攝人臉圖片或人臉模型的非活體人臉影像所採集並標記為負樣本的非活體人臉影像，該人臉採集全域影像包括人臉區域和背景區域，該第一深度學習模型是分類模型，提取該人臉採集全

域影像的結構性特徵，並將人臉採集全域影像分類至活體採集影像類別或者非活體採集影像類別，且其中，該第二深度學習模型包括分別與該人臉區域的鼻部對應的鼻部深度學習模型和與眼部對應的眼部深度學習模型，該人臉區域部分影像包括人臉區域但不包括背景區域，該第二深度學習模型是分類模型，提取該人臉區域部分影像的影像材質性特徵，並將該人臉區域部分影像分類至活體採集影像類別或者非活體採集影像類別，且

其中，該結構性特徵包括至少以下任一種：屏幕照片邊緣以及人臉畸變，且該影像材質性特徵包括至少以下任一種：模糊、紋理及顏色失真。

【第9項】

如申請專利範圍第8項所述的裝置，該第一深度學習模型與該第二深度學習模型採用不同的深度學習網路結構。

【第10項】

如申請專利範圍第9項所述的裝置，該不同的深度學習網路結構的以下至少一種參數不同：隱變量層數、隱變量層類型、神經元節點數、卷積單元的卷積核大小。

【第11項】

如申請專利範圍第8項所述的裝置，該人臉採集全域影像具有第一標籤，該第一標籤表明其對應的人臉採集全域影像是活體採集影像還是非活體採集影像；

該訓練模組根據人臉採集全域影像，訓練第一深度學

習模型，具體包括：

該訓練模組將人臉採集全域影像輸入第一深度學習模型，該第一深度學習模型根據該人臉採集全域影像的影像結構性特徵，預測該人臉採集全域影像是活體採集影像還是非活體採集影像；以及

根據預測結果和該人臉採集全域影像的該第一標籤，對該第一深度學習模型進行調整。

【第12項】

如申請專利範圍第8項所述的裝置，該人臉區域部分影像具有第二標籤，該第二標籤表明其對應的人臉區域部分影像是活體採集影像還是非活體採集影像；

該訓練模組根據從該人臉採集全域影像中裁剪的人臉區域部分影像，訓練第二深度學習模型，具體包括：

該訓練模組獲取從該人臉採集全域影像中裁剪的人臉區域部分影像；

將該獲取的人臉區域部分影像輸入第二深度學習模型，該第二深度學習模型根據該人臉區域部分影像的影像材質性特徵，預測該人臉區域部分影像是活體採集影像還是非活體採集影像；以及

根據預測結果和該人臉區域部分影像的該第二標籤，對該第二深度學習模型進行調整。

【第13項】

如申請專利範圍第8項所述的裝置，該第一深度學習模型與該第二深度學習模型係基於卷積神經網路；以及

該第一深度學習模型所基於的卷積神經網路的卷積單元的卷積核大於該第二深度學習模型所基於的卷積神經網路的卷積單元的卷積核。

【第14項】

如申請專利範圍第8項所述的裝置，該檢測模組根據訓練後的該第一深度學習模型和該第二深度學習模型，進行人臉活體檢測，具體包括：

該檢測模組獲取為進行人臉活體檢測採集的人臉採集全域影像；

將該採集的人臉採集全域影像輸入訓練後的該第一深度學習模型進行處理，得到對應的第一預測資料；

獲取從該採集的人臉採集全域影像中裁剪到人臉區域部分影像，並輸入訓練後的該第二深度學習模型進行處理，得到對應的第二預測資料；以及

透過根據該第一預測資料和該第二預測資料進行聯合決策，得到針對該用戶刷臉影像的人臉活體檢測結果。

【第15項】

一種電子設備，包括：

至少一個處理器；以及，

與該至少一個處理器通訊連接的記憶體，其中，

該記憶體儲存有可被該至少一個處理器執行的指令，該指令被該至少一個處理器執行，以使該至少一個處理器能夠：

根據人臉採集全域影像，藉由對多個全域影像進行監

督式的訓練來訓練第一深度學習模型，其中，該全域影像包括藉由拍攝活體人臉所採集並標記為正樣本的活體人臉影像，以及藉由拍攝人臉圖片或人臉模型的非活體人臉影像所採集並標記為負樣本的非活體人臉影像，該人臉採集全域影像包括人臉區域和背景區域，該第一深度學習模型是分類模型，提取該人臉採集全域影像的結構性特徵，並將人臉採集全域影像分類至活體採集影像類別或者非活體採集影像類別，其中，該結構性特徵包括至少以下任一種：屏幕照片邊緣以及人臉畸變；

根據從該人臉採集全域影像中裁剪的人臉區域部分影像，藉由對多個裁剪的人臉區域部分影像進行監督式的訓練來訓練第二深度學習模型，其中，該第二深度學習模型包括分別與該人臉區域的鼻部對應的鼻部深度學習模型和與眼部對應的眼部深度學習模型，該人臉區域部分影像包括人臉區域但不包括背景區域，該第二深度學習模型是分類模型，提取該人臉區域部分影像的影像材質性特徵，並將該人臉區域部分影像分類至活體採集影像類別或者非活體採集影像類別，其中，該影像材質性特徵包括至少以下任一種：模糊、紋理及顏色失真；以及

根據訓練後的該第一深度學習模型對第一全域影像進行人臉活體檢測以獲得第一預測得分，以及使用多個訓練後的該第二深度學習模型獲取第二預測得分，包括：

獲取採集的第一全域影像用於人臉活體檢測；

將該第一全域影像輸入訓練後的該第一深度學習

模型中進行處理以獲得該第一預測得分；

獲取從該第一全域影像裁剪的多個提取的人臉影像，該提取的人臉影像包括眼部影像區域影像和鼻部影像區域影像，並將該提取的人臉影像輸入到訓練後的包括該眼部深度學習模型和該鼻部深度學習模型的該第二深度學習模型中進行處理，以獲得第二預測得分；

基於該第一預測得分和該第二預測得分，產生預測得分結果；以及

比較該預測得分結果與閾值，以確定該第一全域影像是活體影像還是非活體影像。

【發明圖式】

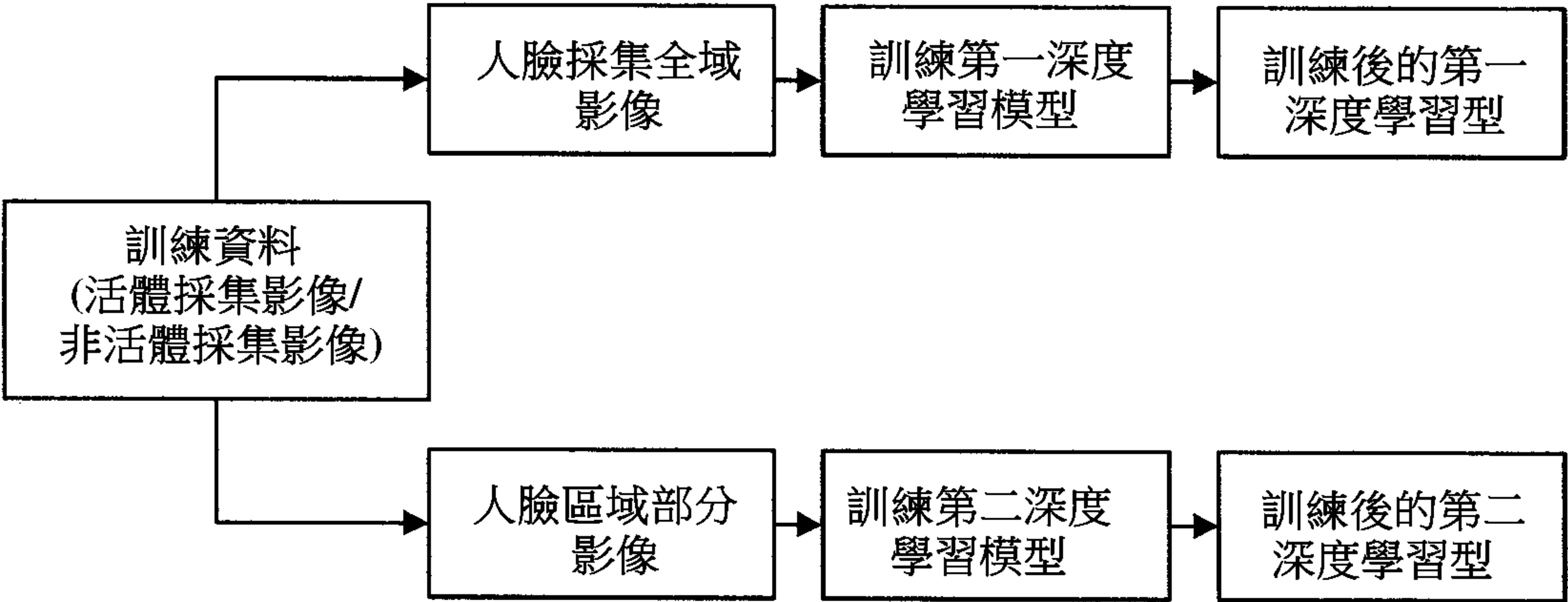


圖 1

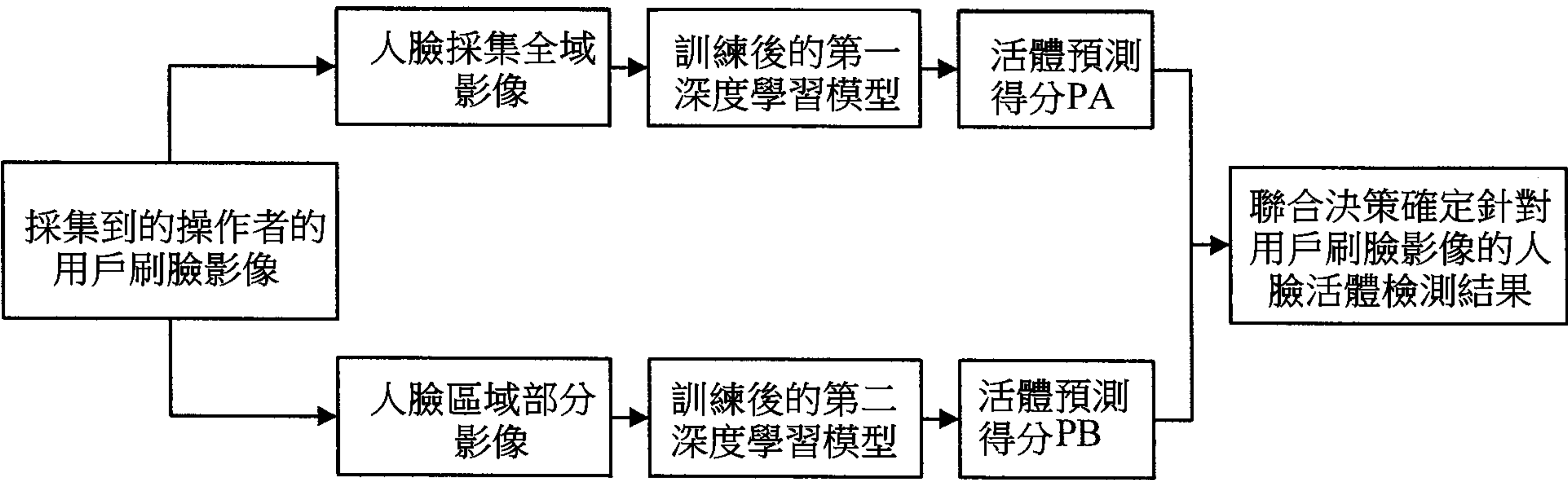


圖 2

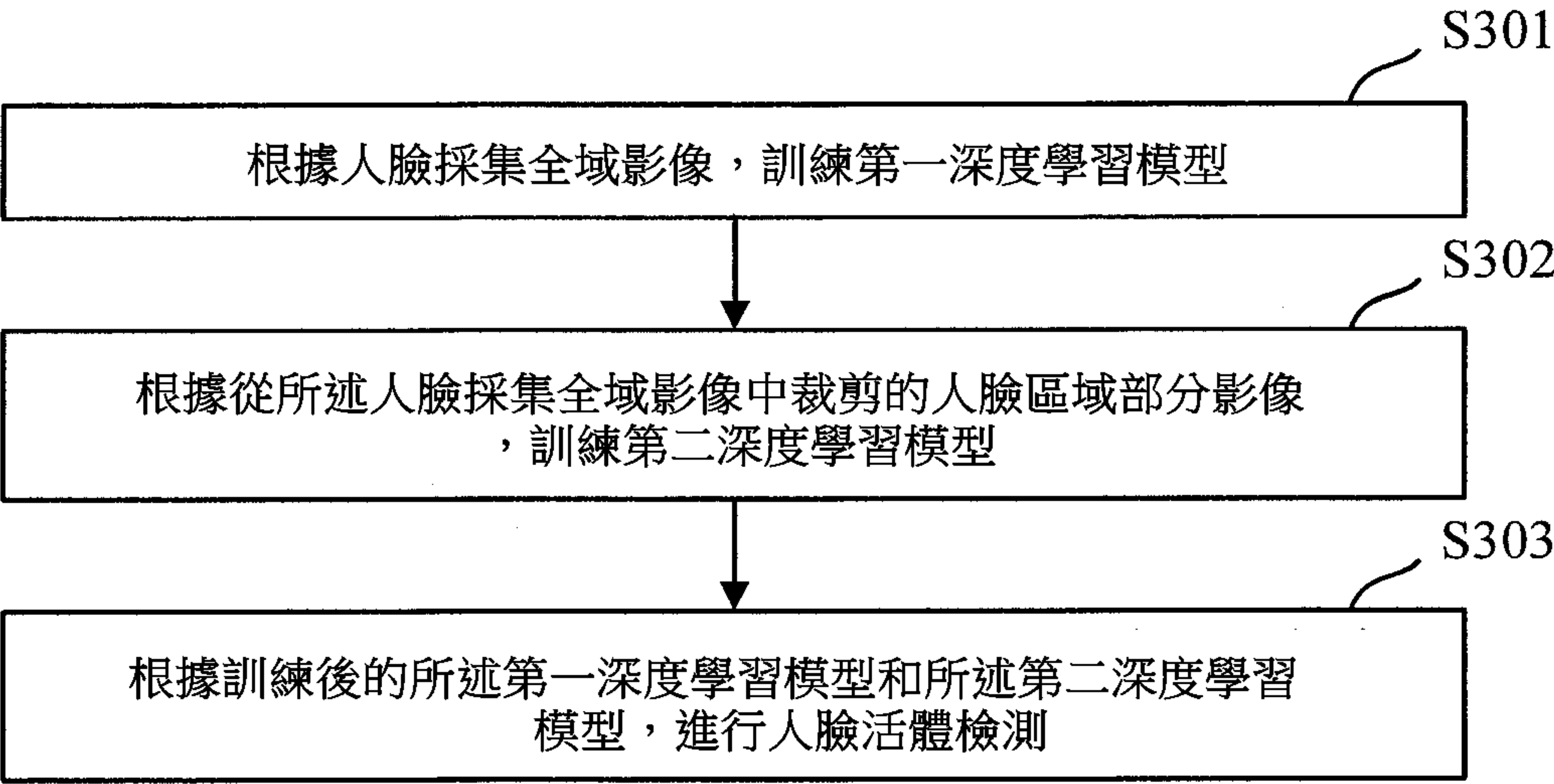


圖 3

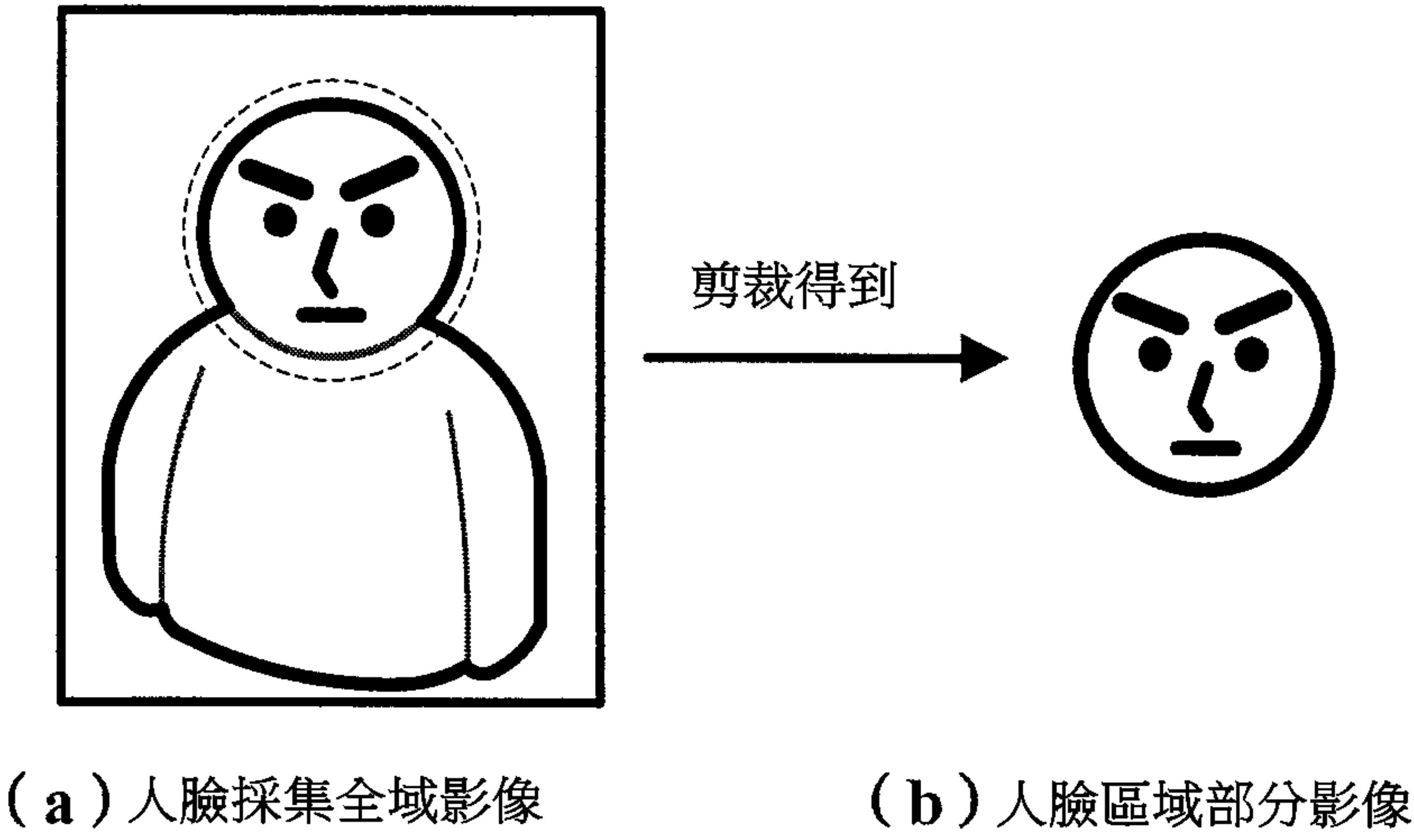


圖 4

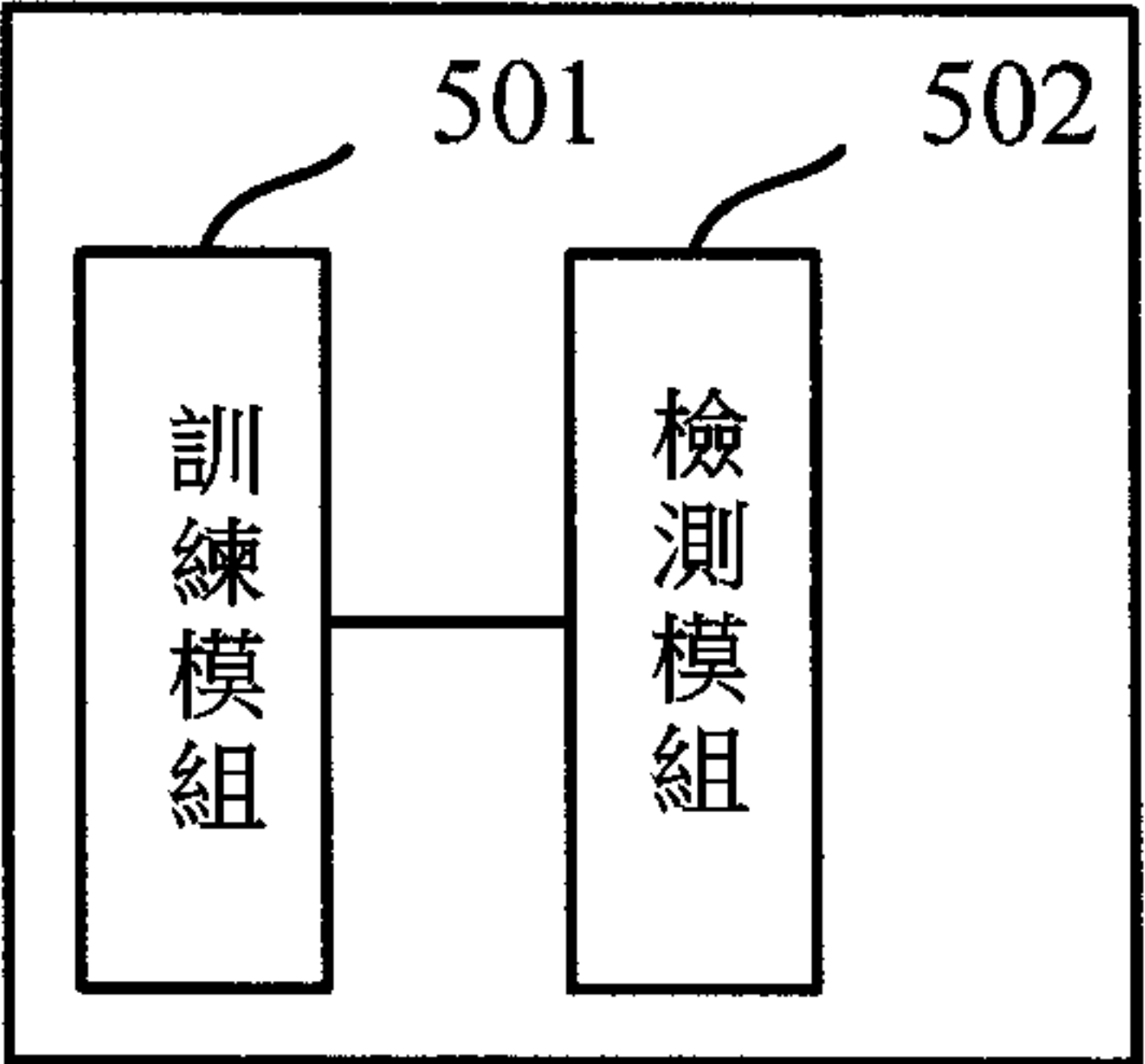


圖 5