

社會隱性：重新思考軌跡

預測評估及有效性

隱式最大似然估計

阿卜杜拉·穆罕默德¹ , Deyao Zhu² , 和沃倫 Vu¹

穆罕默德·埃爾霍賽尼², *克ристи安·克洛德爾^{1,*}

¹ 德州大學奧斯汀分校

{abdullah.mohamed, warren.vu, christian.claudel}@utexas.edu KAUST

²

{deyao.zhu, mohamed.elhoseiny}@kaust.edu.sa

³* 平等建議

抽象的。Best-of-N (BoN) 平均位移誤差 (ADE)/最終位移誤差 (FDE) 是評估軌跡預測模型最常用的指標。然而，BoN 並未量化整個生成的樣本，導致模型的預測品質和性能的視圖不完整。我們提出了一個新的指標，平均馬氏距離 (AMD) 來解決這個問題。AMD 是量化整個生成樣本與真實情況的接近程度的指標。

我們還引入了平均最大特徵值 (AMV) 指標，用於量化預測的總體分佈。我們的指標經過實驗驗證，顯示 ADE/FDE 對分佈變化不敏感，與 AMD/AMV 指標不同，給人一種準確度有偏差的感覺。我們引入使用隱式最大似然估計 (IMLE) 作為傳統生成模型的替代品來訓練我們的模型社會隱性。IMLE 訓練機制與 AMD/AMV 預測軌跡的目標一致，該軌跡接近真實情況且分佈緊密。Social-Implicit 是一種記憶體高效的深度模型，只有 5.8K 參數，在約 580Hz 的即時運行速度並取得了有競爭力的結果。

關鍵字：運動預測、運動預測、深度圖 CNN、評估、軌跡預測

1 簡介

軌跡預測是多種應用的重要組成部分，例如自動駕駛[6,18,5,41,34]、擴增實境[20,40]和機器人[21,3]。

通常，解決這個問題需要一個生成模型來預測未來智能體的軌跡。儘管有許多深層模型和設計架構可以解決這個問題，但所使用的評估方法受到質疑。

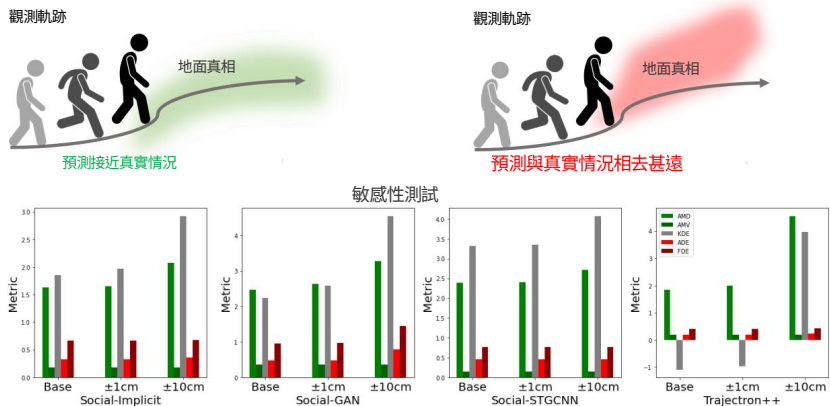


圖 1 :目前的 BoN ADE/FDE 指標對預測分佈不敏感。BoN ADE/FDE 僅關注最接近真實情況的樣本。我們可以看到，對於綠色和紅色預測，BoN ADE/FDE 保持不變。另一方面，提議的 AMD/AMV 指標會根據整個預測分佈的接近程度及其相對於真實情況的分佈而變化。這使得 AMD/AMV 成為評估預測的更好指標。

通常，使用兩個指標來評估軌跡預測模型。

第一個是平均位移誤差 (ADE)[29]，它是預測軌跡和地面真實軌跡之間的平均L2距離。較低的 ADE 值意味著整體預測軌跡接近地面實況。另一個指標是最終位移誤差 (FDE) [1]，它是兩個最終預測位置和地面真實位置之間的L2距離。換句話說，它描述了預測的智能體是否達到了最後的目標。此外，FDE 越低，模型在預測過程中不累積誤差的效果就越好。

在使用基於循環的架構的先前工作中注意到了累積錯誤導致更高 FDE 的問題。先前的工作引入了基於 CNN 的完整架構 [27] 的想法來解決這種錯誤累積行為。

然而，這種 ADE/FDE 指標仍然不適合產生模型。生成模型預測未來軌跡的多個樣本，隱式形成預測分佈。這種生成行為適合該問題，因為代理人或行人的運動可以是具有可能的未來軌跡的多模態。為了在生成設定中使用 ADE/FDE，[1,9] 的工作引入了 Best-of-N (BoN) 的概念。BoN 技術從 N 個樣本中選擇最接近真實情況的樣本，並計算其 ADE/FDE 度量。這有一個主要問題，就是忽略生成的樣本集。

模型可能會產生一個離群樣本，幸運的是，該樣本足夠接近真實情況，而其他樣本則遠離真實情況。這種方法在現實生活中的應用上也失敗了，因為缺乏評估

的預測。一些重要的組件，例如運動規劃和碰撞避免，需要完整的預測視圖。我們注意到的另一個問題是，基於 ADE/FDE 指標的最新模型 [36,43,26,23] 在 ETH [29] 和 UCY 上僅相差 1 厘米 ADE 和幾厘米 FDE [16] 數據集，該領域最常用的資料集之一。前一個 SOTA 模型與下一個 SOTA 模型之間 1 公分的差異是如此微妙和微小，以至於可能是註釋錯誤或異常採樣。

因此，需要一種新的測量來評估整個預測樣本，並了解整個生成的分佈與真實情況的關係。此外，還需要量化生成樣本的不確定性，以提供有關模型置信度的視圖，這是現實應用中所需要的。為此，我們引入了 Maha-lanobis 距離 [24] 作為該領域的測量。我們引入了兩個指標：平均馬氏距離 (AMD)，用於評估產生的分佈與真實情況的接近程度；以及平均最大特徵值 (AMV)，用於評估預測的置信度。AMD 以標準差單位的形式量化了地面實況點與預測分佈的接近程度。此外，AMD 與 χ 分佈相聯繫，幫助我們在已知產生的分佈自由度時確定預測的置信度。AMV 取決於預測分佈的協方差矩陣特徵值的最大幅度。它量化了預測的傳播。因此，我們可以透過使用一個模型來判斷它是否比另一個模型更有信心。因此，我們的目標是實現一個模型，該模型產生的分佈接近真實情況，並且樣本分佈在真實情況周圍較小。這個目標使我們重新思考用於訓練運動預測模型的生成模型的本質。我們可以將所使用的生成技術分為參數生成技術和非參數生成技術。參數模型使用最大似然估計 (MLE) 將預測軌跡建模為高斯或高斯混合模型 (GMM)。

生成對抗網路 (GAN) [8] 是非參數分佈的一個例子。這些方法學習觀察到的軌跡的分佈，以便產生未來的軌跡。然而，軌跡預測模型的主要目標是產生的樣本本身。MLE 需要大量樣本才能收斂，而我們在實務上卻沒有這種能力。而 GAN 則依賴判別器的設計，而 VAE 需要最佳化證據下限 (ELBO)。因此，我們需要一種只專注於生成的樣本並且不會帶來額外麻煩的生成方法。在這項工作中，我們證明隱式最大似然估計 (IMLE) 技術是這些方法的有效替代方法。IMLE 直接關注預測軌跡，簡化了最佳化函數。透過使用 IMLE 來訓練我們引入的模型 Social-Implicit，與先前的工作相比，預測軌跡在品質和準確性方面都有所提高。Social-Implicit 是一種記憶體高效的深度模型，只有 5.8K 個參數，比最接近的 SOTA 少了近 55 倍，並且實時運行速度比最接近的 SOTA 快了近 8.5 倍。

這項工作的組織如下：我們首先回顧最近的相關文獻

作品。然後我們制定運動預測問題，然後介紹和討論新指標 AMD 和 AMV。然後我們介紹

用於訓練我們的社會隱式模型的軌跡條件 IMLE 機制。我們透過解釋社會隱性的架構來遵循這一點。最後，我們分析了我們模型的新指標和最近的 SOTA 指標的結果並附有敏感度分析。

2. 文獻綜述

軌跡預測模型最近的工作提出了各種模型預測未來的軌跡。根據它們的輸出格式，它們可以大致分為兩類。將未來明確建模為參數分佈，或將未來隱式建模為非參數分佈。在

第一類，方法將未來明確建模為連續或離散分佈 [1,27,32,4,36,38,5,19,44,15,31,42,43]。例如，S-LSTM [1] 和 S-STGCNN [27] 使用高斯分佈來建模未來軌跡。透過最大似然估計（MLE）進行訓練。高斯分佈是單一模式，無法趕上未來的多模態。為了解決這個問題，問題，PRECOG [32]、Trajectron++ [36]、ExpertTraj [43] 和 AgentFormer [42] 學習潛在行為分佈，可以是離散的[36,43]或連續的[32,42]，以表示代理的多模態意圖。在這些作品中，預測的高斯分佈是根據採樣的潛在樣本產生的意圖。這類方法通常是基於條件 VAE [37]。除了連續分佈，如 MTP [5] 和 LaneGCN [19] 等高斯方法使用離散分佈來表示未來。這些方法預測一個固定的作為未來候選的確定性軌跡的數量，並使用分類分佈來模擬它們的可能性。第二類，一些方法以隱式方式對未來分佈進行建模。例如，S-GAN [9]、So-Phie [35]、S-BiGAT [13] 和 DiversityGAN [10] 遵循條件 GAN [7] 建築學。它們不是產生分佈作為模型輸出，而是預測以隨機取樣雜訊為條件的確定性軌跡並透過對抗性損失機制進行訓練。我們提出的方法 Social-Implicit 透過使用 IMLE 進行訓練來隱式地對未來分佈進行建模 [17] 避免額外的麻煩，例如 GAN 訓練機制中的判別器。

軌跡預測指標 大多數軌跡預測方法是透過平均位移誤差 (ADE) [29] 或最終位移誤差 (FDE) [1] 指標進行評估。這兩個指標是基於 L2 距離整個時間範圍 (ADE) 或預測和地面真實軌跡之間的最後時間步長 (FDE)。當模型生成的分佈為輸出，Best-of-N 技巧 [9] 僅用於評估最佳軌跡來自 N 個採樣的預測。平均 ADE/FDE 也可用於評估的預測，它最適合單模態預測以及當預測接近高斯分佈。在多模態中，當例如，預測彼此矛盾（向左轉、向左轉）

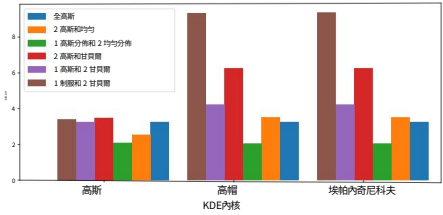
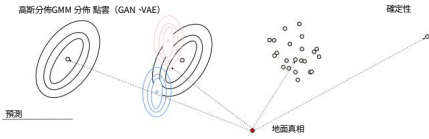


圖 2：不同運動預測模型的輸出。高斯和 GMM 是參數模型的例子。GAN 和 VAE 是圖 3 的範例：

不同的混合模型 NLL 非參數模型。最後一個

cate-KDE 與 KDE Ker-gory 的選擇是確定性模型輸出。內爾。越低越好。我們也不需要統一的指標來評估結果因這些模型而異的所有情況。內核。

右)平均 ADE/FDE 將會失敗，因為它本質上是確定性的。評估分佈品質的另一種方法是核密度估計 (KDE)，首次在[12]中使用。KDE 根據預測樣本擬合基於核的分佈，並將真實情況的負對數似然估計為評估分數。奎爾等人。[30]提出了一種綜合度量，它是不同相似性度量的加權和，以減輕度量偏差。但他們的指標僅適用於確定性模型。我們提出了兩個新指標平均馬哈拉諾比斯距離 (AMD) 和平均最大特徵值 (AMV)，它們是 BoN ADE/FDE 評估預測的更好替代方案。

3 平均馬氏距離 (AMD) 指標

我們將軌跡預測問題定義如下：給定 N 個智能體在一系列觀察時間步 T_0 上的觀察軌跡，目標是預測下一個 T_p 預測時間步。觀察到的軌跡包含 P 個點，其中每個點表示智能體的空間位置。在行人軌跡預測問題中， P 是二維笛卡爾位置 (x, y) 。我們將觀察集表示為 $do = \{pt | t \in T_0\}$ 且預測集為 $dp = \{pt | t \in T_p\}$ 。

為了強調目前 BoN ADE/FDE 中的問題，我們從圖 2 開始，它說明了不同類型的預測模型輸出。對於確定性模型，計算公式 1 中定義的 ADE/FDE 指標非常簡單。但對於產生模型，ADE/FDE 是透過 BoN 方法計算的。BoN 的工作原理是採樣 N 個（通常是 20 個）樣本，選擇最接近真實情況的樣本，然後使用樣本計算 ADE/FDE。我們可以從多個方面批評這種 BoN 方法。主要問題是它不會量化整個生成的樣本，而只關注最接近的樣本。與密度完全偏離地面事實但有一個接近地面的樣本的另一個模型相比，這可能會對密度圍繞地面事實的模型不利。

真相。我們可以在預告圖 1 中看到這一點。我們的另一個擔憂是，使用這種 BoN 方法，可以運行該指標幾次，得到的結果比另一個模型好 1 公分。在某些極端情況下，幸運的隨機運行可能具有非常低的 BoN ADE/FDE。[12] 的工作注意到了這個問題，並引入了方程式 1 中定義的核密度估計 (KDE) 的使用。KDE 是一個基於核的工具，可以獲取預測機率密度的非參數表示。然後，計算真實值的負對數似然度並以對數單位 (nats) 報告。

然而，KDE 指標存在多種限制和擔憂。主要的問題是，KDE 指標在樣本數量較少的情況下對核心的選擇很敏感，這就是現實資料集中的情況。圖 3 說明了 KDE 中使用的內核的不同選擇與各種發行版的混合。我們注意到，當使用高斯核時，它不區分不同的樣本，並且與其他輸出相比可能更傾向於具有完整 GMM 輸出的模型。我們也注意到，每當使用不同的內核選擇時，例如使用高帽內核與高斯內核，我們可能會得到混合的結果。[12] 的工作是將 KDE 指標與高斯內核結合使用。KDE 核心的另一個限制是它不包含易於解釋的分析屬性。此限制是由於 KDE 的非參數性質造成的。這些令人感興趣的屬性可能是機率矩和置信區間。

(1)

其中 p 是智能體 $n \in N$ 在預測時間步 $t \in T_p$ 處的真实位置，且 p_n 是預測位置。新的度量需要是參數化的，允許進一步分析並且對其計算距離的方式不敏感。

因此，我們介紹馬氏距離的用法。馬哈拉諾比斯距離可以測量點距分佈的距離，同時將此距離與預測變異數相關聯。它還具有將其與卡方分佈聯繫起來的分析屬性，其中可以評估預測的置信度。最後，它取決於高斯分佈，這使得可以進一步分析預測矩。馬哈拉諾比斯距離 (MD) 定義為： $\sum -1(p - \mu)$ 。其中， μ 是預測的平均值，

$$MD_{\mu, \Sigma, p} = (p - \mu)^T \Sigma^{-1} (p - \mu)$$

預測分佈的方差， p 是地面真實位置。

最初，馬氏距離並不是為 GMM 分佈設計的。然而，[39] 的工作將 MD 擴展為 GMM，並將其表示為：

(2)

其中 G 是每個混合成分的逆協方差，按機率平均和加權，定義為： $G =$

$$\frac{\sum_{k=1}^K \pi_k \int_{\mathbb{R}^D} p(x|k) dx}{\sum_{k=1}^K \pi_k \int_{\mathbb{R}^D} p(x|k) dx}, \text{ 在哪裡}$$

K 是混合成分的數量， π_k 是第 k 個成分的權重，GMM 的平均值定義為： $\mu_{GMM} = \sum_{k=1}^K \pi_k \mu_k$ 。
積分

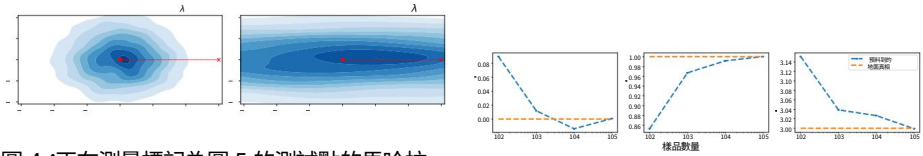


圖 4：正在測量標記為圖 5 的測試點的馬哈拉

諾比斯距離：GMM 擬合顯示誤差和 x 。顯示了兩個雙變量高斯分佈 - 需要大量樣本才能收斂到分佈，左邊的一個是真實平均值，某些東西可能比右邊的非參數模型具有更低的方差。 λ 代表最大 els。在 1000 個樣本標記處，分佈的 GMM 絕對特徵值開始收斂到真實平均值和協方差。與穩定 MD 的變異數。

G^A 中的項是易於處理的，如 [39] 所述。我們注意到，如果 GMM 僅包含一個分量，則 G 將是 $\Sigma^A - 1$ ，因此 Tipping 的 MD 是原始 MD 的更廣義的版本。我們的方法如下，無論模型產生什麼分佈或輸出，我們都適合 GMM。將會提出關於最佳混合分量 K 的數量的問題。這可以透過使用貝葉斯資訊準則 (BIC) 輕鬆解決： $BIC = m \ln n - 2 \ln L$ GMM，其中 m 是 GMM 模型的參數數量， n 是觀測資料點的數量， L GMM 是模型的似然函數。BIC 越低，表示資料點的擬合 GMM 模型越好。根據 BIC 自動選擇最佳 GMM。研究從已經預測軌跡平均值和變異數 (例如 [1, 27]) 的模型中進行取樣的原因，我們希望公平。擬合 GMM 會產生一定的誤差，因此我們希望將這種誤差納入所有測量模式中，以具有統一的測量。圖 5 顯示了這個錯誤。

因為確定性模型沒有方差，所以我們需要模型中誤差的表示。我們可以多次訓練確定性模型並將預測擬合到 GMM。另一個建議是計算整體平均值和方差，並直接應用 MD 距離，而不進行 GMM 擬合。後者方法可能會產生與 GMM 擬合誤差相當的誤差，使指標更加公平。在補充資料中，我們討論了這兩種情況。

我們認為，評估確定性模型與生成模型是一個需要進一步研究的懸而未決的問題，它是一個類似於 KDE [12] 限制的限制。現在，我們定義平均馬氏距離 (AMD)：

(3)

4 平均最大特徵值 (AMV) 指標

AMD 指標的一個主要問題是它與分佈的變異數高度相關。模型可以預測未來的軌跡

實際情況存在巨大差異，真實情況接近平均值。與具有較高方差的另一個模型相比，這將導致 AMD 非常低。另一個例子是一個模型，它預測覆蓋所有預測點的巨大變異數（以公尺為單位）。這也將導致最佳 AMD 值。為了對抗這種錯誤行為，我們需要我們的模型具有低 AMD 和低方差，也就是更確定的模型。此外，在實際應用中，我們需要量化預測的整體不確定性，以便對效能有整體的了解。因此，我們介紹了協方差矩陣特徵值的用法。協方差矩陣的最大幅值特性值是協方差矩陣散佈的指標。圖 4 顯示了兩種分佈，左邊的變異數比右邊的變異數小。我們注意到，與右分佈相比，固定點相對於左分佈的 MD 要高得多。然而，左側分佈的最大幅值特徵值遠小於右側分佈，顯示了預測的擴散。因此，為了正確評估模型，我們需要 AMD 和價差測量。正如我們所討論的，我們可以直接從預測協方差矩陣中測量分佈。由於我們在 AMD 指標中引入的框架，我們有一個預測的協方差矩陣。KDE 指標中缺少一些東西。現在，我們可以介紹 AMV 指標：

(4)

其中 λ_1 是矩陣特徵值中幅度最大的特徵值。 Σ GMM 是預測 GMM 分佈的協方差矩陣，定義為： Σ GMM = 。因此，AMV 成為評估預測軌跡整體分佈的指標。AMD 較低的模型的預測分佈將更接近真實情況。AMV 較低的模型的預測會更確定。因此，與具有較高 AMD/AMV 平均值的另一個模型相比，具有較低 AMD/AMV 平均值的模型是首選。為此，我們使用 AMD+AMV 作為良好模型的指標。

2

5 軌跡條件隱式最大似然估計 (IMLE) 機制

從 AMD/AMV 指標的目標出發，我們需要一個能夠產生接近真實情況且擴散程度較低的樣本的模型，我們需要一種訓練機制，以允許完全控制預測樣本作為主要最佳化目標。典型的訓練機制，例如最大似然估計 (MLE) 或其變體，例如最大化證據下限 (ELBO)，鼓勵預測樣本接近某些（地面真實）資料樣本。這樣，可能會遺漏一些數據範例並導致模式

下降[17]。其他方法，例如 GAN，需要引入額外的模組，例如判別器，它們的訓練通常不穩定，需要仔細調整以達到適當的納許均衡。[17]的工作引入了隱式最大似然估計 (IMLE) 的概念。IMLE 鼓勵每個目標地面實況接近某些預測樣本。因此，與 MLE 不同，它會導致預測分佈更好地覆蓋真實情況。IMLE 透過簡單的機制訓練模型：將雜訊注入模型的輸入以預測多個樣本，選擇最接近真實情況的樣本，然後使用該樣本進行反向傳播。與其他生成方法不同，IMLE 不使用特定的訓練技術來加載優化目標，並且由於簡單的基於距離最小化的優化而保持訓練穩定。使用 IMLE 作為訓練機制與 AMD/AMV 目標保持一致，並專注於重要的產品，即預測輸出。IMLE 的另一個觀點是，與由製程和測量雜訊驅動模型的卡爾曼濾波器等估計技術相比，它是一種更先進的神經技術。我們建議讀者參考原始 IMLE 論文 [17] 進行進一步討論。訓練機制如 Alg.1 所示：

演算法 1 軌跡條件隱含最大似然估計 (IMLE) 演算法需求：資料集 $D = (d)$ 以輸入為條件 要求：選擇適當的損失函數 $L(.)$ ，例如如均

方误差或 L_1 $\sum_{i=1}^n$ 以及有取樣機制的模型 $\theta(.)$

將模型初始化為 $e = 1$

到 Epochs do 從 D 選取隨機

批次 (d_o, d_p) 繪製 iid 樣本 $d \sim p$

$\sigma(i) \leftarrow \arg \min_{d \sim p} L(d - d_o) \quad \forall i$

$\sigma(i) \leftarrow \arg \min_{d \sim p} L(d - d_o) \quad \forall i$

$\theta \leftarrow \theta - \eta \nabla \theta \sigma(i)$

回傳 θ 結

束

6 社會隱性模型

在本節中，我們介紹社會隱式模型。Social-Implicit 記憶體很小，只有 5.8K 參數，即時運行時間為 588Hz。此方法包括三個概念：社交區域、社交細胞和社交損失。

社交區域：社交區域根據觀察到的代理軌跡的最大速度變化對它們進行聚類。行人平均速率為 1.2m/s [14]。我們注意到我們可以將行人的移動分為四組。第一組是靜止組，以行人在紅綠燈等待為例。該組的最大速度變化在 0-0.01m/s 之間。雖然第二組是運動量最小的行人，

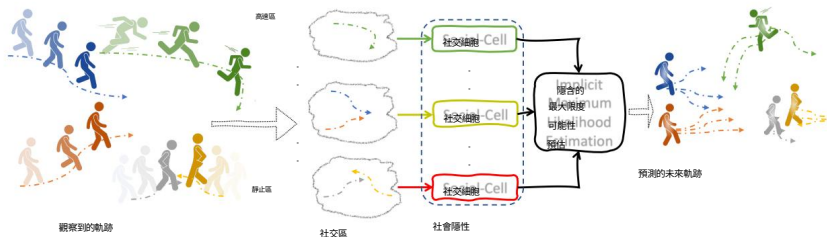


圖 6 :社會隱式模型概念。社交區域根據觀察到的最大速度對觀察到的軌跡進行聚類。然後每個社交區域由社交單元處理。此模型使用 IMLE 進行訓練。

又稱原地搖晃的人或互相打招呼的一群行人，通常該群體的最大速度變化在 0.01-0.1m/s 之間。

第三組是平均步行速度的行人，這些行人的移動在 0.1-1.2m/s 之間。最後一組是奔跑的行人，平均速度通常高於 1.2m/s。當深度模型針對靜止的行人和速度更快的行人進行訓練時，預測中將存在對移動行人的偏見。這將迫使模型將不移動的物體預測為移動的物體。這是一種資料不平衡，或者換句話說，是零（靜止）膨脹的資料問題。因此，需要社交區域的概念來解決這個問題。根據經驗，我們表明，帶有社交區域的模型比沒有社交區域的模型表現更好。Social-Zones 的輸入是觀察到的軌跡，輸出是行人簇，每個簇是一個維度為 $P \times T_o \times N$ 的圖。

社交細胞：社交隱式模型的基本建構單元是社交細胞。Social-Cell 是一個 4 層深度模型，簡單且直接處理觀測的時空圖。圖 7 說明了社交單元的結構。我們注意到該單元有兩個元件，一個元件在本地層級處理每個單獨的代理，另一個元件在全域層級處理整個代理。我們透過自學習權重結合本地和全域流來產生單元的最終輸出。局部流和全域流都是兩個連續的殘差連接 CNN 層。第一個 CNN 是空間 CNN，它創建觀察代理的空間資訊的嵌入。第二層是時間 CNN，處理觀察到的軌跡的時間方面。它將時間視為特徵通道，讓我們在不使用循環網路的情況下預測下一個 T_p 時間步長 [27]。

我們發現這種簡單的架構與更大、更複雜的模型一樣有效，從而產生較小的記憶體大小和即時運行時間功能。

每個社交單元處理一個特定的社交區域。輸入為 $P \times T_o \times N$ ，輸出為 $P \times T_p \times N$ 。操作如圖 7 所示。社會損失：社會隱式的損失函數分為幾個部分。第一部分是我們之前討論過的 IMLE 機制的直接最佳化目標。

第二部分是三重態損失。這個三元組損失認為錨點是

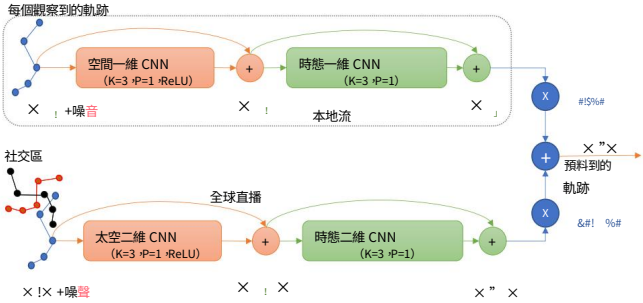


圖 7：社交細胞模型。本地流和全域流只有兩個 CNN。P 是觀測位置，To和Tp是觀測和預測的時間步長，N是代理的數量。其中 CNN 的 K 和 P 是內核和填充尺寸。

最接近的樣本 $\sim d_p$ 真相。正例是下一個最接近的範例 $\sim d_{2-p}$ 真相。反面例子 $\sim d_m$ 來自地面真相。這有助於將樣本分組到更靠近地面的位置事實，導致真實軌跡周圍的分佈更緊密。最後一部分損失的幾何損失函數將預測位置視為多邊形。首先，它確保預測位置之間的內部距離匹配地面實況位置之間的內部距離。其次，它使得確保預測點之間的角度與地面實況點之間的角度相同。它確保預測的場景在幾何上

看起來像是基本事實。我們在公式 5 中定義了這些損失。場景的各個方面都可以超出我們介紹的範圍來解決開放研究領域。

$$\begin{aligned} & \text{---} \\ & \text{---} \end{aligned} \tag{5}$$

$$\begin{aligned} & \text{---} \\ & \text{---} \end{aligned}$$

因此，我們將社會損失定義為：

$$\tag{6}$$

其中 $\alpha_1 = 0.0001$ 、 $\alpha_2 = 0.0001$ 或 0.00001 、 $\alpha_3 = 0.0001$ 。

7 實驗與分析

我們從整體表現和敏感度分析分析了常見行人運動預測模型的指標行為。然後我們從設計組件和性能方面分析我們的模型。

7.1 指標敏感性分析與評估 我們表明，BoN ADE/FDE 指標對分

佈的變化或偏移不敏感，而 AMD 和 KDE 可以量化這種變化。圖 1 說明了這個概念。我們透過使用不同的量（特別是 $\pm 1\text{cm}$ 和 $\pm 10\text{cm}$ ）移動預測樣本來測試不同的模型。在所有模型中，BoN ADE/FDE 指標完全沒有變化，或有非常微小的細微變化。與測量整個分佈的指標（如 AMD 和 KDE）不同，預測分佈的變化反映在指標中。由此得出的結論是，ADE/FDE 度量對整個分佈的變化並不敏感，即使是 10cm 的巨大變化，有時可以定義一個新的 SOTA 模型而不是另一個模型。因此，BoN ADE/FDE 指標無法評估整個預測軌跡。此外，AMV 指標保持不變，這是預期的，因為僅改變預測不會改變變異數。我們注意到 Trajectron++ 的 KDE 與其他模型不同，因為 Trajectron++ 輸出是 GMM 分佈，正如我們之前討論的那樣，由於內核選擇，KDE 度量存在偏差。

為了定量評估指標，我們使用 ETH/UCY 資料集報告不同運動預測模型的 AMD/AMV、KDE 和 ADE/FDE 指標。我們選擇了經典的，例如 S-GAN [9] 和 S-STGCNN [27]。我們也選擇了更新的版本，例如 Trajectron++ [36] 和 ExpertTraj [43]。從選項卡。在圖 1 中，我們注意到最後兩個模型（被認為是 SOTA）在 ADE/FDE 指標上有幾分點的差異。然而，當我們使用 AMD/AMV 指標評估它們時，我們注意到 Trajectron++ 的效能比 ExpertTraj 模型好得多。根據 AMD/AMV 指標，Expert-Traj 產生一個不圍繞真實情況的緊密分佈，這會導致與 Trajectron++ 不同的更高的 AMD。儘管 ExpertTraj 和 Trajectron++ 的 ADE/FDE 指標非常接近，但整個預測樣本的品質卻完全不同。檢查我們的模式「社交隱式」的結果，我們發現它的 AMD/AMV 最低。透過消化結果，ADE/FDE 指標並不表示與上述敏感性分析相關的模型的整體表現。我們在史丹佛無人機資料集（SDD）[33] 上測試我們的模型和指標。我們遵循 SOTA 模型 DAG-Net [28] 的設定。表 2 的實驗結果顯示我們的模型優於 DAG-Net。這與 ETH/UCY 資料集的結果一致。

7.2 社會隱含的消融研究

我們對社會隱性成分進行了消融研究。具體來說，就是社交區域和社交損失。標籤。圖 3 示出了結果。我們注意到

表 1 :對於所有指標 ,越低越好 。結果在 ETH/UCY 上數據集 。M 是未報道的型號 。NaN 表示計算失敗 。專家軌跡 ADE/FDE 取自他們的論文 。有時我們會注意到 ,即使是一個模型 ADE/FDE 較低 ,AMD/AMV 透過評估整體表現與此相矛盾產生的樣本 。

	以末坊	飯店	大學	札拉1	札拉2	平均值 (AMD+AMV)/2	
	ADE/FDE AMD/AMV 在每種						
S-GAN [9]	0.81/1.52 3.94/0.373 5.02	0.72/1.61 2.59/0.384 3.45	0.60/1.26 2.37/0.440 2.03	0.34/0.69 1.79/0.355 0.68	0.42/0.84 1.66/0.254 -0.03	0.58/1.18 2.47/0.361 2.23	1.42
S-STGCNN [27]	0.64/1.11 3.73/0.094 6.83	0.49/0.85 1.67/0.297 1.56	0.44/0.79 3.31/0.060 5.65	0.34/0.53 1.65/0.149 1.40	0.30/0.48 1.57/0.103 1.17	0.44/0.75 2.39/0.104 3.32	1.26
Trajectron++ [36]	0.39/0.83 3.04/0.286 1.34	0.12/0.21 1.98/0.114 -1.89	0.20/0.44 1.73/0.228 -1.08	0.15/0.33 1.21/0.171 -1.38	0.11/0.25 1.23/0.116 -2.43	0.19/0.41 1.84/0.183 -1.09	1.01
專家軌跡 [43]	0.30/0.62 61.77/0.034 南	0.09/0.15 21.30/0.003 南	0.19/0.44 毫米 中號	0.15/0.31 32.14/0.005 南	0.12/0.24 毫米 中號	0.17/0.35 38.4/0.004 南	19.20
社會隱性 (我們的)	0.66/1.44 3.05/0.127 5.08	0.20/0.36 0.58/0.410 0.59	0.31/0.60 1.65/0.148 1.67	0.25/0.50 1.72/0.078 1.24	0.22/0.43 1.16/0.106 0.69	0.33/0.67 1.63/0.174 1.85	0.90

表 2 :對於所有指標 ,越低越好 。結果位於 SDD 資料集上 。

	ADE FDE AMD AMV KDE (AMD+AMV)/2	
STGAT [11]	0.58 1.11 -	- - -
社交方式 [2]	0.62 1.16 -	- - -
DAG網路[28]	0.53 1.04 3.17 0.247 1.76	1.70
社會隱性 (我們的)	0.47 0.89 2.83 0.077 3.89	1.45

社交區域的存在使 AMD 指標提高了近 40% 。
它還帶來了良好的 AMV 值 ,從而增強了 AMD/AMV 的整體性能 。我們注意到 ,僅社交區域領先的三重態損失就有所改善
AMD/AMV 。幾何角度損失對AMD/AMV的改善作用大於幾何距離損失 。雖然兩者一起工作效果更好 。

表 3 :社會隱性成分的消融研究 。

L-triplet LG 距離 LG 角度區域 AMD/KDE/AMV AMD/AMV ADE/FDE					
				2.06/2.29/0.110 1.09	0.32/0.66
			✓	2.04/1.86/0.104 1.07	0.32/0.64
✓			✓	1.96/2.16/0.097 1.03	0.56/1.08
✓	✓		✓	2.13/2.32/0.092 1.11 1.03	0.32/0.62
✓	✓	✓	✓	2.13/2.32/0.092 1.11 1.84/1.78920.	0.78/1.38
✓	✓	✓	✓	1.16 1.63/1.85/0.174 0.90	0.35/0.71
✓	✓	✓	✓		0.33/0.67

7.3 推理與記憶分析

與其他模型相比 ,社會隱式模型除了是最準確的模型之外
AMD/AMV 指標上的模型 ,它是最小且最快的

參數大小和推理時間。表 4 顯示了這些結果。最近的 SOTA 是 ExpertTraj，它的 Social-Implicit 體積小 55 倍，速度快 8.5 倍。

7.4 社交區消融

標籤。圖5顯示了區域數的消融。不同區域影響模型的性能。此外，該模型對區域速度很敏感。例如，當我們將最後一個區域從 1.2m/s 更改為 0.6m/s 時，結果發生了變化。1.2m/s反映了人類平均步行速度，因此0.6m/s不適合數據，因此與 1.2m/s 相比，它導致性能較差。

表 4：參數計數和報告的平均推理速度，以 GTX1080Ti 為基準。

	參數	計數	速度 (s)
S-GAN [9]	46.3K	(7.98x)	0.0968 (56.9x)
S-STGCNN [27]	7.6K	(1.3x)	0.0020 (1.2x)
Trajectron++ [36]	128K	(22.1x)	0.6044 (355.5x)
專家軌跡 [43]	323.3k	(55.74x)	0.0144 (8.5x)
社會隱式 (我們的) 5.8K			0.0017

表 5：區域數量及其速度對我們模型精度的影響。

#Zones	ADE/FDE	KDE	AMD/AMV	AVG
1	0.35/0.71	2.76	2.29/0.090	1.16
2	0.36/0.73	3.09	2.32/0.088	1.20
3	0.34/0.70	2.31	2.08/0.085	1.08
4@0.6m/s	0.34/0.69	2.85	2.31/0.080	1.19
4@1.2m/s	0.33/0.67	1.85	1.63/0.174	0.90

7.5 定性結果

在圖 8 中，我們列出了我們的方法和基線模型的兩個定性範例。在第一行，我們看到一個行人在地面真實未來的盡頭右轉。我們注意到 Social-Implicit 和 Trajectron++ 涵蓋了真實的未來，而 S-GAN 和 ExpertTraj 為我們提供了早期轉向預測，集中註意力遠離基本事實。第二排呈現之字形行走行人。S-STGCNN、Trajectron++ 和 ExpertTraj 等基線模型無法產生良好的分佈來覆蓋地面真實軌跡，這與我們的和 S-GAN。雖然ExpertTraj的預測接近地面事實上，ExpertTraj 過於自信，與事實真相相矛盾。定性結果顯示我們的預測分佈較好。我們也展示了多代理圖 9 中的交互作用。ExpertTraj 過於自信而忽略了真實情況，S-STGCNN 與碰撞有很大的方差，Trajectron++ 具有真實情況接近預測的分佈尾部，而我們的則具有適當的平衡。更多的定性結果見附錄。

8 結論

我們引入了 AMD 和 AMV 指標，用於評估軌跡預測模型產生的分佈。我們證明了 BoN ADE/FDE 指標對生成的分佈給出了不充分的評估。基於 AMD/AMV 指標的目標是建立一個產生樣本的模式

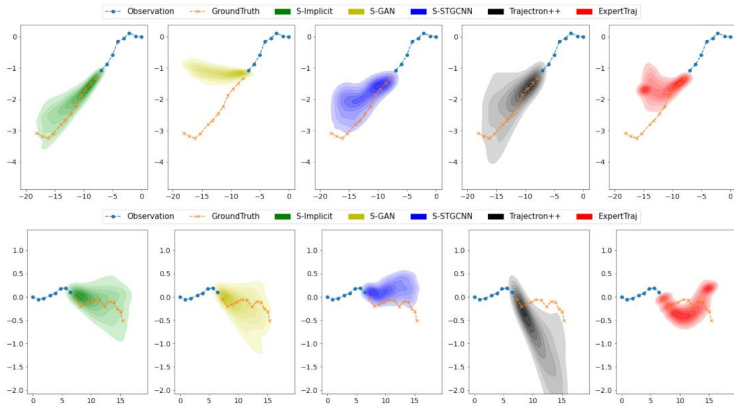


圖 8 :ETH/UCY 資料集上預測軌跡的可視化。

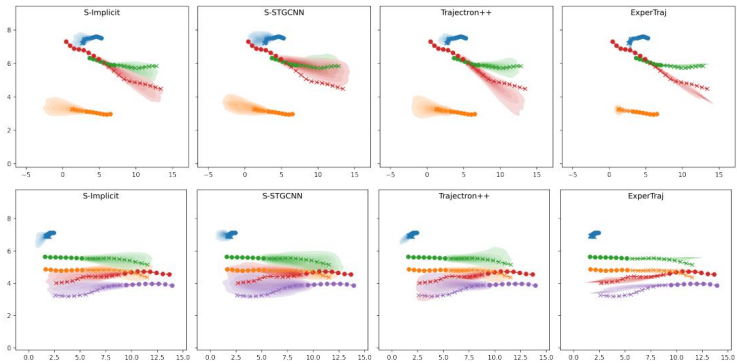


圖 9 :ETH/UCY 資料集上的多人交互案例。

由於接近真實情況且變異數很小，我們引入了使用 IMLE 來訓練我們的模型 Social-Implicit。我們證明了社交隱式是一種即時運行的高效記憶模型，並依賴社交區域、社交細胞和社交損失等多個概念來提高表現。

總的來說，我們邀請運動預測社群調整 AMD/AMV，以便更好地評估他們的方法。

參考

1. Alahi, A. \Goel, K. \Ramanathan, V. \Robicquet, A. \Fei-Fei, L. \Savarese, S. :社會
Istm :擁擠空間中的人體軌跡預測。請參閱 :IEEE 會議錄
電腦視覺和模式識別會議。第 961–971 頁 (2016)
2. Amirian, J. \Hayet, JB \Pettre, J. :社交方式 :使用甘斯學習行人軌跡的多態態分佈。請參閱 :IEEE/CVF 計算機
視覺和模式識別研討會會議記錄。第 0–0 頁 (2019)
3. Bùutepage, J. \Kjellström, H. \Kragic, D. :預測許多未來 :線上人類
人機互動的運動預測與生成。請參閱 :2018 年 IEEE 機器人與自動化國際會議 (ICRA) 。第 4563–4570 頁。
IEEE
(2018)
4. Chai, Y. \Sapp, B. \Bansal, M. \Anguelov, D. :多路徑 :用於行為預測的多個機率錨軌跡假設。 arXiv 預印本
arXiv:1910.05449 (2019)
5. Cui, H., Radosavljevic, V., Chou, FC, Lin, TH, Nguyen, T., Huang, TK, Schnei-der, J., Djuric, N. :使用多態態
軌跡預測自動駕駛
深度卷積網路。於 :2019年機器人與技術國際會議
自動化 (ICRA) 。第 2090–2096 頁。 IEEE (2019)
6. Deo, N., Trivedi, MM :用於車輛軌跡預測的捲積社會池。請參閱 :IEEE 電腦視覺與模式會議論文集
認可研討會。第 1468–1476 頁 (2018)
7. Gauthier, J. :用於卷積人臉生成的條件生成對抗網路。史丹佛大學 CS231N 課程計畫 :卷積神經網絡
視覺識別。2014年冬季學期(5), 2 (2014)
8. Goodfellow, I. \Pouget-Abadie, J. \Mirza, M. \Xu, B. \Warde-Farley, D. \Ozair, S. ,
Courville, A. \Bengio, Y. :生成對抗網路。神經資訊處理系統的進展 27 (2014)
9. Gupta, A. \Johnson, J. \Fei-Fei, L. \Savarese, S. \Alahi, A. :社交甘 :社交
生成對抗網路的可接受軌跡。在 :會議記錄
IEEE 電腦視覺和模式識別會議。第 2255 頁–
2264 (2018)
10. 黃X. \麥吉爾, SG \德卡斯特羅, JA \弗萊徹, L. \倫納德, JJ \威廉斯, BC \Rosman, G. :Diversitygan :透過潛在語意採樣進行多樣性感知車輛運動預測。 IEEE 機器人與自動化快報
5(4), 5089–5096
(2020)
11. 黃宇 \畢宏 \李志 \毛濤 \王志 :Stgat :時空建模
人類軌跡預測的交互作用。請參閱 :IEEE/CVF 會議記錄
電腦視覺國際會議。第 6272–6281 頁 (2019)
12. Ivanovic, B., Pavone, M. :軌跡 :使用動態時空圖進行機率多智能體軌跡建模。請參閱 :IEEE/CVF 計算機視覺國際
會議論文集。第 2375–2384 頁 (2019)
13. Kosaraju, V. \Sadeghian, A. \Martí n-Martí n, R. \Reid, I. \Rezatofighi, SH,
Savarese, S. :Social-bigat :使用 bike-gan 進行多模式軌跡預測
圖注意力網。 arXiv 預印本 arXiv:1907.03395 (2019)
14. Laplante, JN, Kaeser, TP :行人步行速度的持續演變
假設。交通工程師學會。ITE 期刊 74(9), 32 (2004)
15. Lee, N. \Choi, W. \Vernaza, P. \Choy, CB \Torr, PH \Chandraker, M. :願望 :
與交互代理的動態場景中的遙遠未來預測。請參閱 :IEEE 計算機視覺和模式識別會議論文集。頁數
336–345 (2017)

16. Lerner, A. \Chrysanthou, Y. \Lischinski, D. :以人群為例。在：計算機圖形論壇。卷。 26，第 655–664 頁。Wiley 線上圖書館 (2007)
17. Li, K., Malik, J. :隱式最大似然估計。arXiv 預印本 arXiv:1809.09087 (2018)
18. Li, X., Ying, X., Chuah, MC :Grip :基於圖形的交互感知軌跡預言。請參閱 :2019 年 IEEE 智慧交通系統會議 (ITSC)。頁數3960–3966。IEEE (2019)
19. Liang, M., Yang, B., Hu, R., Chen, Y., Liao, R., Feng, S., Urtasun, R.: Learning 用於運動預測的車道圖表示。參見 :歐洲會議電腦視覺。第 541–556 頁。施普林格 (2020)
20. Limmer, M., Forster, J., Baudach, D., Schüle, F., Schweiger, R., Lensch, HP :用於擴增實境導航系統的基於穩健深度學習的道路預測晚上。請參閱 :2016 年 IEEE 第 19 屆智慧交通系統國際會議 (ITSC)。第 1888–1895 頁。IEEE (2016)
21. Liu, H., Wang, L. :人機協作的人體運動預測。製造系統雜誌 44, 287–294 (2017)
22. Liu, L., Wei, W., Chow, KH, Loper, M., Gursoy, E., Truex, S., Wu, Y. :深度神經網絡整合對抗欺騙 :整合多樣性、準確性和穩健性。請參閱 :2019 年 IEEE 第 16 屆移動自組織與感測器系統國際會議 (大量的)。第 274–282 頁。IEEE (2019)
23. Liu, Y., Yan, Q., Alahi, A. :社會性 :社會意識運動表徵的對比學習。請參閱 :IEEE/CVF 國際會議論文集關於電腦視覺。第 15118–15129 頁 (2021)
24. Mahalanobis, PC :論統計學中的廣義距離。國家研究所印度科學 (1936)
25. Makansi, O. \Cicek, O. \Marrakchi, Y. \Brox, T. :揭露交通參與者未來預測中具有挑戰性的長尾。arXiv 預印本 arXiv:2103.12474 (2021)
26. Mangalam, K. \An, Y. \Girase, H. \Malik, J. :從目標、航點和路徑到長期人類軌跡預測。在 :過程。國際會議電腦視覺 (ICCV) (2021 年 10 月)
27. Mohamed, A. \Qian, K. \Elhoseiny, M. \Caudel, C. :Social-stgcnn :用於人類軌跡預測的社會時空圖卷積神經網絡。在 :IEEE/CVF 電腦視覺與模式辨識會議論文集。第 14424–14432 頁 (2020)
28. Monti, A. \Bertugli, A. \Calderara, S. \Cucchiara, R. :Dag-net :雙重關注用於軌跡預測的圖神經網絡。請參閱 :2020 年第 25 屆國際模式識別會議 (ICPR)。第 2551–2558 頁。IEEE (2021)
29. Pellegrini, S. \Ess, A. \Schindler, K. \Van Gool, L. :你永遠不會獨行 :為多目標追蹤建模社交行為。於 :2009 年第12屆IEEE國際會議電腦視覺會議。第 261–268 頁。IEEE (2009)
30. Quehl, J. \Hu, H. \Ta, s, OS, . \Rehder, E. \Lauer, M. :我的預測有多好 ?尋找軌跡預測評估的相似性測量。時間 :2017年 IEEE 第 20 屆智慧交通系統國際會議 (ITSC)。頁數1–6。IEEE (2017)
31. Rhinehart, N. \Kitani, KM \Vernaza, P. :R2p2 :重新參數化的前推多樣化、精確的生成路徑預測的策略。請參閱 :歐洲電腦視覺會議 (ECCV) 會議記錄。第 772–788 頁 (2018)
32. Rhinehart, N. \McAllister, R. \Kitani, K. \Levine, S. :Precog :以視覺多智能體設定中的目標為條件的預測。請參閱 :IEEE/CVF 會議記錄電腦視覺國際會議。第 2821–2830 頁 (2019)

33. Robicquet, A. ·Sadeghian, A. ·Alahi, A. ·Savarese, S. :學習社交禮儀 ·擁擠場景中的人類軌跡理解 ·請參閱 ·歐洲電腦視覺會議 ·第 549–565 頁 ·施普林格 (2016)
34. Rudenko, A ·Palmieri, L ·Herman, M ·Kitani, KM ·Gavrila, DM ·Arras, KO :人體運動軌跡預測 :一項調查 ·國際機器人研究雜誌 39(8), 895–935 (2020) ·
- 35.Sadeghian, A. ·Kosaraju, V. ·Sadeghian, A. ·Hirose, N. ·Rezatofighi, H. ·Savarese, S. :Sophie :一個專注的 gan ·用於預測符合社會和物理約束的路徑 ·請參閱 ·IEEE/CVF 計算機視覺和模式識別會議論文集 ·第 1349–1358 頁 (2019)
36. Salzmann, T. ·Ivanovic, B. ·Chakravarty, P. ·Pavone, M. :Trajectron++ :利用異質資料進行動態可行軌跡預測 ·請參閱 ·電腦視覺 - ECCV 2020 :第 16 屆歐洲會議 ·英國格拉斯哥 ·2020 年 8 月 23 - 28 日 ·會議記錄 ·第 XVIII 部分 16 ·第 683 - 700 頁 ·施普林格 (2020)
37. Sohn, K. ·Lee, H. ·Yan, X. :使用深度條件產生模型學習結構化輸出表示 ·神經資訊處理系統的進展 28, 3483–3491 (2015)
38. Tang, C., Salakhutdinov, RR :多重期貨預測 ·神經科學的進展 資訊處理系統 32, 15424–15434 (2019)
39. Tipping ·ME :從高斯混合模型導出聚類分析距離函數 ·參見 :1999年第九屆人工神經網路國際會議 ICANN 99. (會議出版物第 470 號) ·卷 ·2 ·第 815–820 頁 ·工程技術學院 (1999)
40. Westphal, C. :支援擴增實境和虛擬的網路挑戰 現實 · IEEE ICNC (2017)
41. Wu, P., Chen, S., Metaxas, DN :Motionnet :基於鳥瞰圖的自動駕駛聯合感知和運動預測 ·請參閱 ·IEEE/CVF 計算機視覺和模式識別會議論文集 ·第 11385–11395 頁 (2020)
42. Yuan, Y., Weng, X., Ou, Y., Kitani, K. :Agentformer :用於社會時間多智能體預測的智能體感知變壓器 · arXiv 預印本 arXiv:2103.14023 (2021)
- 43.Zhao, H. ·Wildes, RP :你們要去哪裡 ?具有專家目標範例的動態軌跡預測 ·請參閱 ·IEEE/CVF 國際電腦視覺會議論文集 ·第 7629–7638 頁 (2021)
44. Zhu, D., Zahran, M., Li, LE, Elhoseiny, M. :連續空間中不可能訓練的運動預測 ·請參閱 ·第五屆機器人學習年會 (2021)

互動演示

我們引入了一個互動式演示 ·該演示顯示了當生成的分佈發生變化或移動時 ADE ·FDE ·AMD 和 AMV 的變化 ·示範網址 : <https://www.abduallahmohamed.com/social-implicit-amdamv-ade-fde-demo> ·

透過使用此演示 ·人們可以看到更改分佈的直接影響以及 ADE/FDE 指標如何不足以評估預測品質 ·例如 ·當 x 或 y 方向之一的偏移很大時 ·ADE/FDE 將保持恆定 ·

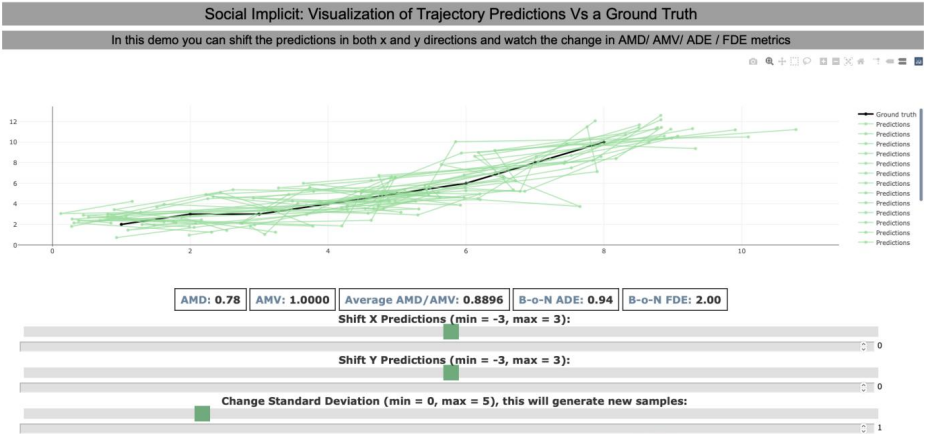


圖 10：社交隱式互動演示。此演示顯示了產生的分佈的指標變化。

B 定性分析

圖 12 和 13 顯示了我們的模型表現良好或可能表現不佳的情況。從圖12開始，在第一排和第二排，我們看到一個行人在過去左轉，在未來直行。第一種情況下的S-GAN和第二種情況下的S-GAN、S-STGCNN、ExpertTraj自信地認為行人將來會右轉。但行人實際上是筆直前進的，我們的方法和Trajectron++ 正確預測了這一點。在第三種情況中，S-GAN 和 S-STGCNN 給我們的預測太慢，而ExpertTraj 給我們的預測太快且顛覆。相較之下，我們的方法和 Trajectron++ 的預測分佈很好地涵蓋了地面事實。在最後一種情況下，ExpertTraj 透過將預測的集中度放在真實值上而表現良好。我們根據觀察到的運動的原始趨勢做出了錯誤的預測。第二張圖13的第一排和第二排，行人中間有一個突然轉彎的未來。

儘管所有方法都無法預測這種轉向，但我們的方法產生的預測分佈最好地涵蓋了真實情況。第三行顯示行人一動也不動。所有方法都為我們提供了接近無運動的預測。其中ExpertTraj的動作是最小的。我們的是第二小的。最後一排顯示一個行人直行，但在未來的中間改變了車道。我們的方法和Trajectron++ 很好地覆蓋了地面真實軌跡，而 S-STGCNN 錯過了新車道，而 ExpertTraj 產生了一個不存在的轉彎。總體而言，儘管 ADE/FDE 指標表明 Trajectron++ 和 ExpertTraj 是最先進的方法，但我們展示了幾個案例，顯示密度遠離真實情況。因此，與量化整個生成的分佈的 AMD/AMV 指標不同，ADE/FDE 無法充分了解模型的準確性。這與結果相關。

表 [1] 和我們對實驗部分的分析，其中我們的模型在 ADM/AMV 指標上表現最佳。

C 確定性模型的評估

我們在 ETH/UCY 資料集上訓練 Social-STGCNN [27] 作為確定性模型。它不是預測高斯分佈，而是預測軌跡目錄。訓練使用MSE作為損失函數。我們想要測試評估確定性模型的兩個假設。第一個是對其進行多次訓練，並使用整合來尋找每個預測軌跡的平均值和變異數。另一個透過計算平均值和變異數將前一個相加，但擬合 GMM，然後從此 GMM 進行取樣。在這個實驗中，我們使用不同的隨機種子對 Social-STGCNN 進行了 3 次訓練。結果如表6所示。報告了第一個設定的AMD和AMV。KDE 不是因為沒有方法可以在不採樣的情況下根據均值和方差計算它，這與我們的度量 AMD 不同，AMD 通過直接將均值和方差插入馬哈拉諾比斯距離方程來具有這種能力。對於第二個設置，當我們將樣本擬合到 GMM 擬合中，然後對多個樣本進行取樣時，報告了 AMD、AMV 和 KDE。我們只使用了 Social-STGCNN 的 3 個整合來模擬現實生活中的情況，也就是說，對其進行 1000 次訓練並從中創建一個整合是不可行的。我們在表 6 中註意到，第二個設定表現出非常大的 AMD 和 KDE，這表明 GMM 擬合沒有收斂，因為我們只有 3 個樣本。通常，我們使用 1000 個樣本來確保 GMM 收斂，因此使用第二個設定是不可行的，因為我們需要那麼多樣本才能很好地擬合 GMM 模型。

我們注意到，在第一個和第二個設定中，AMV 值是相同的。

這是預料之中的，因為 AMV 指標是價差的指標。對於第一個設置，AMD 值對於確定性模型來說似乎是合理的，因為[25]的工作表明大多數運動預測問題可以使用線性卡爾曼濾波器來解決。AMV 指標的巨大值也支持了這一點，因為確定性模型沒有那麼大的分佈。我們將其與 ExpertTraj 主要論文中的結果聯繫起來，其中 AMV 值與我們訓練的確定性模型處於同一數量級。換句話說，由於分佈緊密，ExpertTraj 確實表現為確定性模型。我們可以在圖12 和圖 13 中報告的一些視覺案例中註意到這一點。為了進一步分析，我們在圖11 中繪製了從確定性模型集合中產生的一些樣本以及分佈。我們注意到有時預測的分佈可能是接近真實情況，如左上角的範例所示。此外，它也可以完全關閉，就像其他範例一樣。因此，我們認為使用確定性模型的幾個版本的集合是使用 AMD/AMV 指標評估其性能的好方法。此外，以 AMV 指標作為最佳化目標，可以使用有助於鼓勵多樣性的方法來訓練整合 [22]。

數據集	合奏				GMM 擬合				一般的
	AMD	AMV	哪裡	AMD	AMV	哪裡	ADE	FDE	
以太坊	1.45	35.7	飯店	0.36	-	-	27.89	35.7	12.89
19.6 大學	0.62	170.1	-	zara1	-	-	44.26	19.5	11.27
1.18 28.0 zara2	1.03	96.9	-	-	-	-	24.62	169.9	13.30
			-	-	-	-	28.95	28.0	13.86
			-	-	-	-	12.54	96.8	8.36
平均	0.93	70.06	-	-	-	-	27.65	70.0	11.94
							1.43	2.59	

表 6 :確定性案例實驗。我們將 Social-STGCNN [27] 訓練為使用不同隨機種子的確定性模型。第一個設定報告 AMD/AMV 使用整體的平均值和變異數。第二次設定使用 GMM 擬合平均值和變異數來報告 AMD/AMV/KDE 合奏。ADE/FDE 是整體的平均值。

D 社會隱式實施細節

如前所述，社會隱性包括四個區域。表 7 顯示更多有關區域的詳細資訊。我們注意到每個區域使用不同的配置用於產生樣本的隨機雜訊。慢速區域使用噪音與速度更快的區域相比，方差要低得多。我們還顯示了圖層的詳細信息表 8 中的 Social-Cell。本地和全球流共享相同的設計，只不過本地流使用Conv1D，全域流使用Conv2D。我們將雜訊、全域和局部權重初始化為零。噪音權重為乘以隨機分佈產生的樣本，再相加到輸入張量。模型以學習率訓練了 50 個 epoch = 1，則 45 個 epoch 後學習率降至 0.1。批量大小設定為 128。我們使用 SGD 作為最佳化器。我們也使用了增強技術類似於[36]的軌跡來對抗資料集中的一些不平衡。我們用了隨機旋轉幾度，反轉軌跡，翻轉 x,y 位置，將位置抖動一個小值，增加場景中的節點數量。透過將其與另一個場景結合併改變行人的速度。所附程式碼中提供了模型和增強的實作。

區域	速度範圍	噪音
2	0-0.01 m/s	$N(0, 0.052^2)$, 如果是 $eth N(0, 0.1752^2)$
3	0.01-0.1 m/s	$N(0, 1^2)$ 如果 $eth N(1.5^2)$
4	0.1-1.2 m/s	$N(0, 4^2)$
	1.2 毫秒	$N(0, 8^2)$

表 7 :社交區域配置。速度範圍決定是否觀察到軌跡將在區域內或不在區域內。隨機噪音表現出不同的因區域而異。

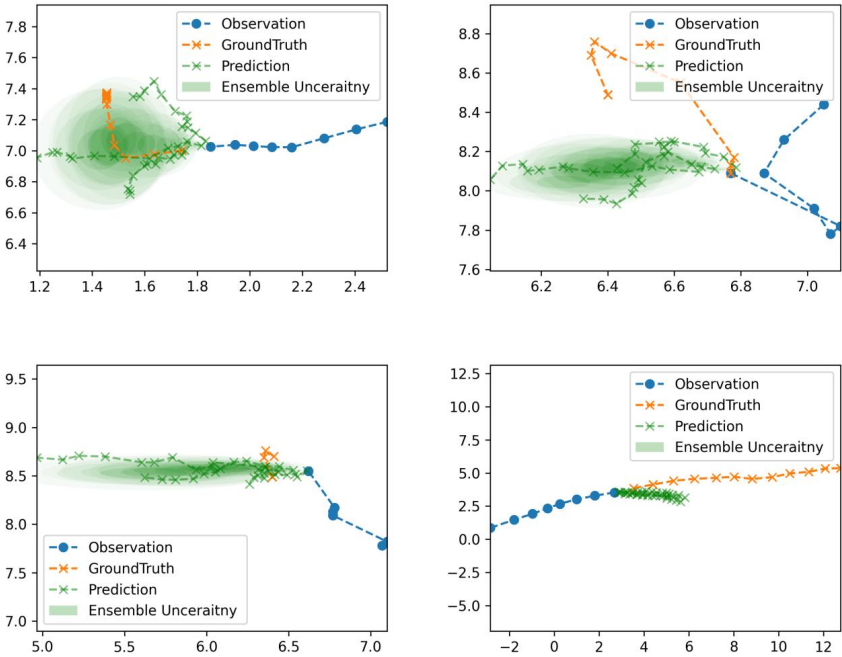


圖 11 :Social-STGCNN 確定性版本預測。

部分	圖層名稱	配置
本地流	空間CNN	Conv1D[P,P,3,1]
	空間啟動ReLU	
	空間 ResCNN	Conv1D[P,P,1,0]
	時態 CNN	Conv1D[To,Tp,3,1]
	Temporal ResCNN	Conv1D[To,Tp,1,0]
全球直播	噪音權重	1 參數
	空間CNN	Conv2D[P,P,3,1]
	空間啟動ReLU	
	空間 ResCNN	Conv2D[P,P,1,0]
	時態 CNN	Conv2D[To,Tp,3,1]
	Temporal ResCNN	Conv2D[To,Tp,1,0]
	全局權重	1 參數
	局部重量	1 參數

表 8 :社交單元配置。Conv1D 或 Conv2D ,其中 $[x,x,x,x] = [\text{輸入特徵}, \text{輸出特徵}, \text{內核大小}, \text{填充大小}]$ 。Res = 加到前一層輸出的剩餘連接。P 是觀察到的維度地點。To和Tp是觀察到的和預測的時間步數。

圖 12 :ETH/UCY 資料集上多個模型預測軌跡的可視化。

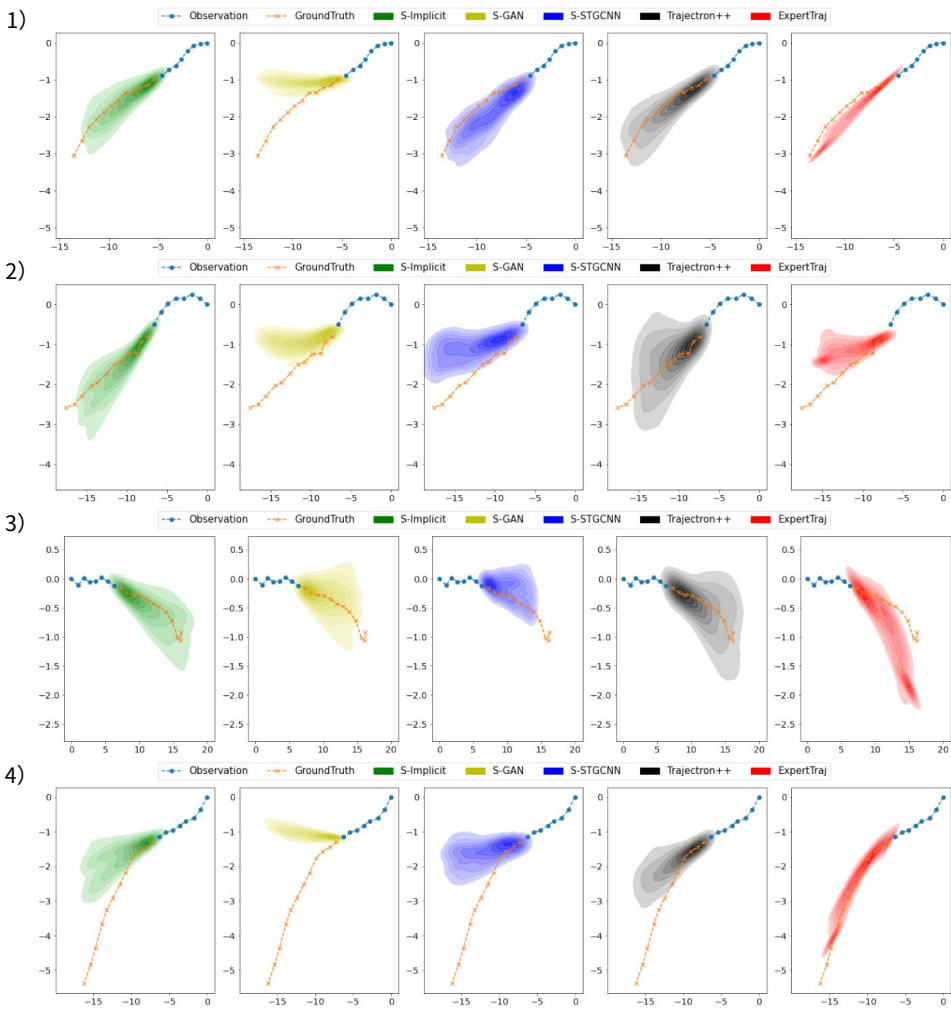


圖 13 :ETH/UCY 資料集上多個模型預測軌跡的可視化。

