

ÜÇ BOYUTLU İSKELET VERİLERİNDEN METRİK ÖĞRENME TABANLI HAREKET TANIMA



Boğaz'da Yapay Öğrenme İsmail Arı Yaz Okulu, 2-5 Temmuz 2018

Şeyma YÜCER, Prof.Dr.Yusuf Sinan AKGÜL

VisLab, Gebze Teknik Üniversitesi, Bilgisayar Mühendisliği Bölümü, Kocaeli, Türkiye

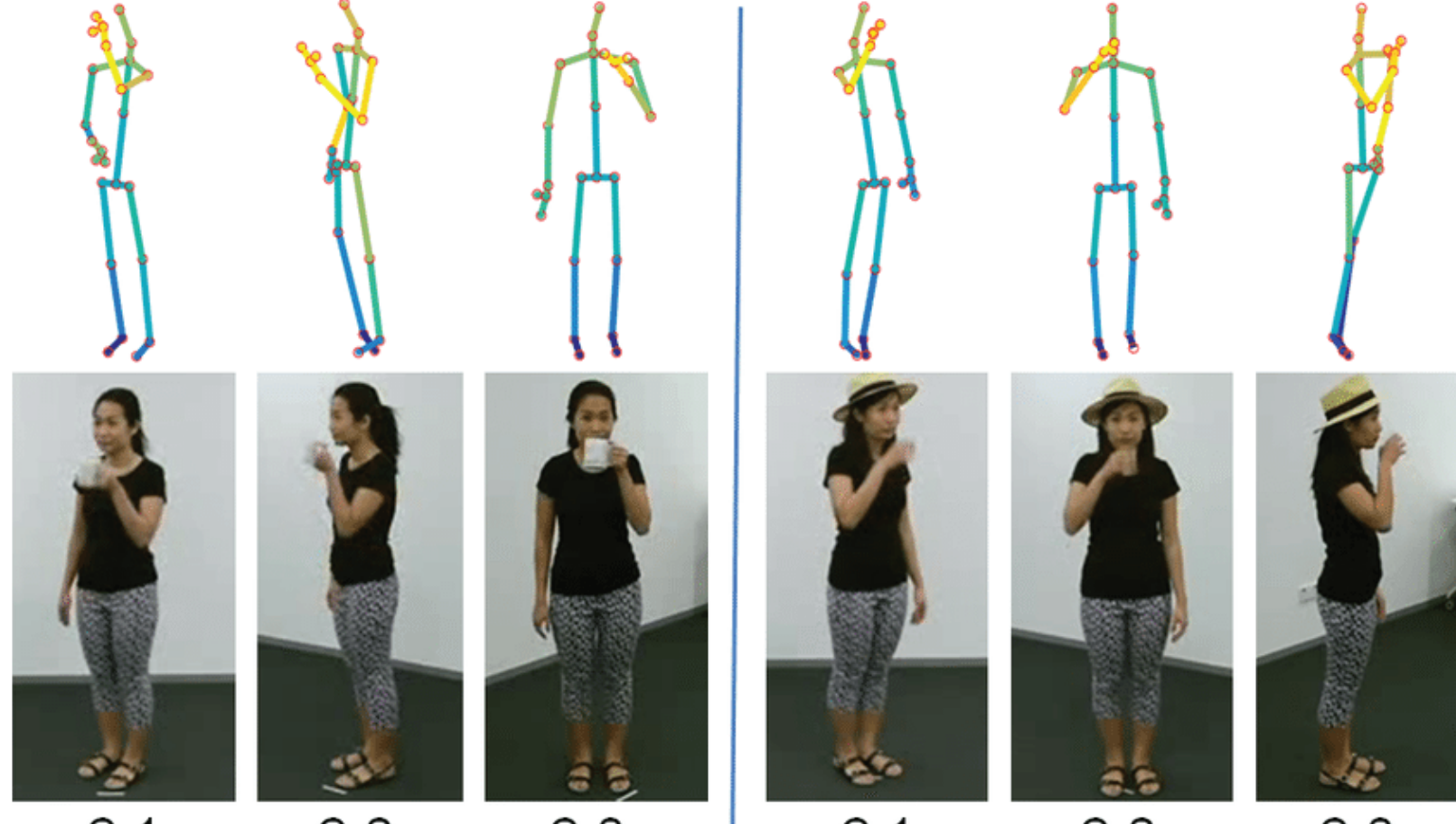
E-posta: {syucer, akgul}@gtu.edu.tr

Özet

Bu çalışmada, 3B iskelet görüntüleri kullanılarak insan hareketlerinin analizi için iki ayrı yöntem önerilmiştir. Bu yöntemlerden ilki, hareketler için gösterim tabanlı bir çözüm sunmaktadır.

Önerdiğimiz ikinci yöntem ise, derin ağ tabanlıdır. Tasarladığımız İkiz LSTM-DML ağı, eylemlerin birbiri ile olan ilişkisini öğrenerek hareketleri tanımlamaktadır. Ağ iki farklı hareketi girdi olarak almaktadır. Her bir hareketin ağıdaki LSTM alt ağları sayesinde zamansal öznitelikleri çıkarılmaktadır. Çift yönlü olarak parametre paylaşımı ile eğitilen ağ, hareketlerin derin metriklerini öğrenmektedir. Böylece baştan sona çalışan ağ, derin metrikleri kullanarak hareketleri sınıflandırabilmektedir. Sınıflandırma dışında hareketlerin benzerliklerini çıkarabilen yöntem diğer çalışmalara kıyasla daha genelleştirilebilirdir.

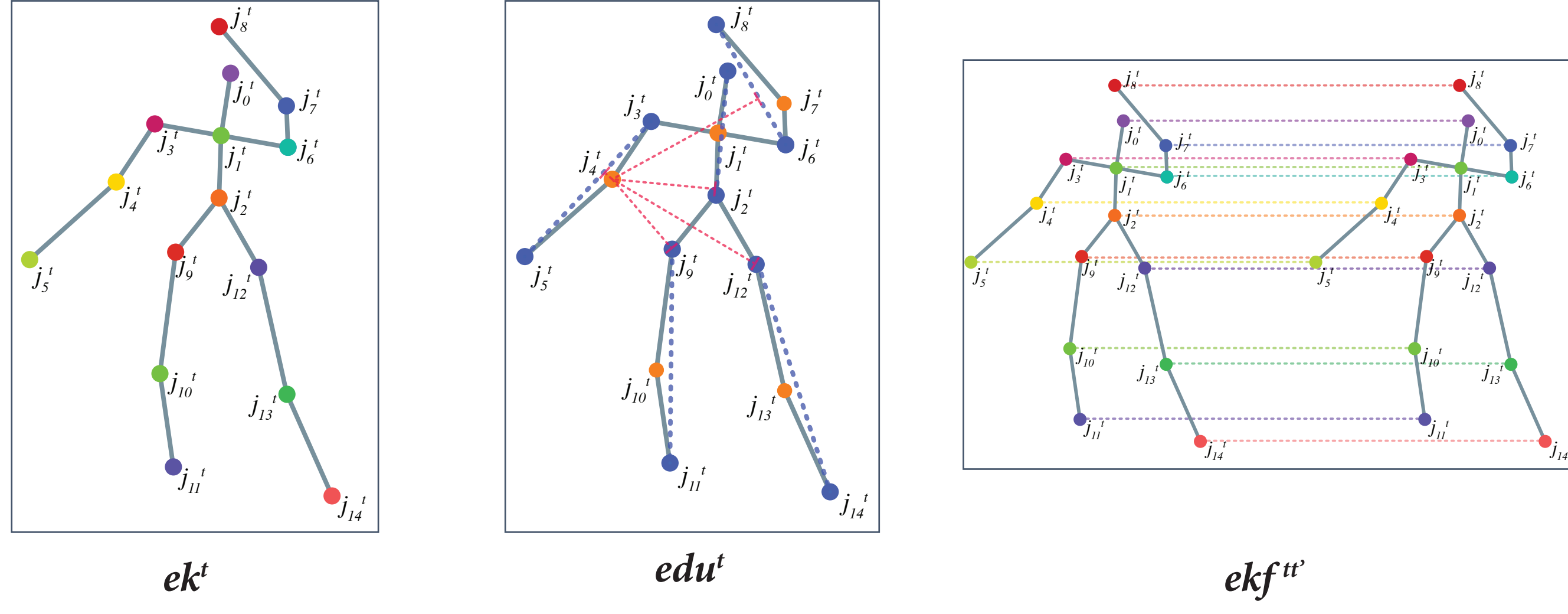
Bu tez için iç mekânda spor ve gündelik hareketlerden oluşan GTU Action 3D veri kümesi oluşturulmuştur. Yöntemlerimiz kendi veri kümemize ek olarak Florence Action 3D, Microsoft, NTU RGB+D veri kümeleri üzerinde test edilmiş ve literatür çalışmaları ile karşılaştırılmıştır.



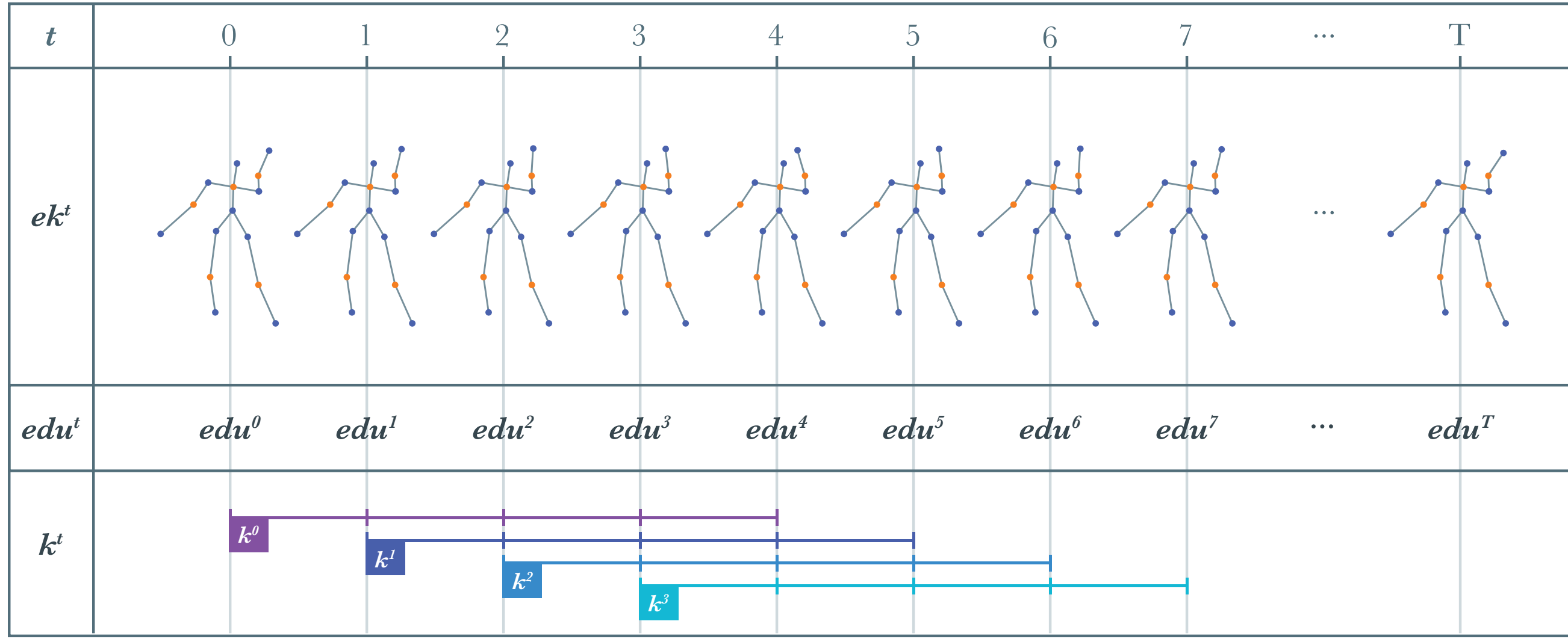
Şekil I. 3B İskelet verileri ve 2 boyutlu görüntüler. [1]

Geometrik Eklem Çantası

Bu kısımda, iskelet görüntülerine ait gösterim oluşturma amacıyla çeşitli geometrik öznitelik performansları denenmiş ve başarımları yüksek öznitelikler çıkarılmıştır. Geometrik öznitelikler, iskelet yapısına ait uzamsal bilgiler ve vücut bölgelerindeki ilişkisel bilgileri içermektedir. Hareketlerden çıkarılan öznitelikler zaman dilimlerine göre parçalanmış ve kelime çantası yöntemi için kelimeler haline getirilmiştir. Kelimelerin histogram vektörleri hesaplanıp sınıflandırılmıştır.

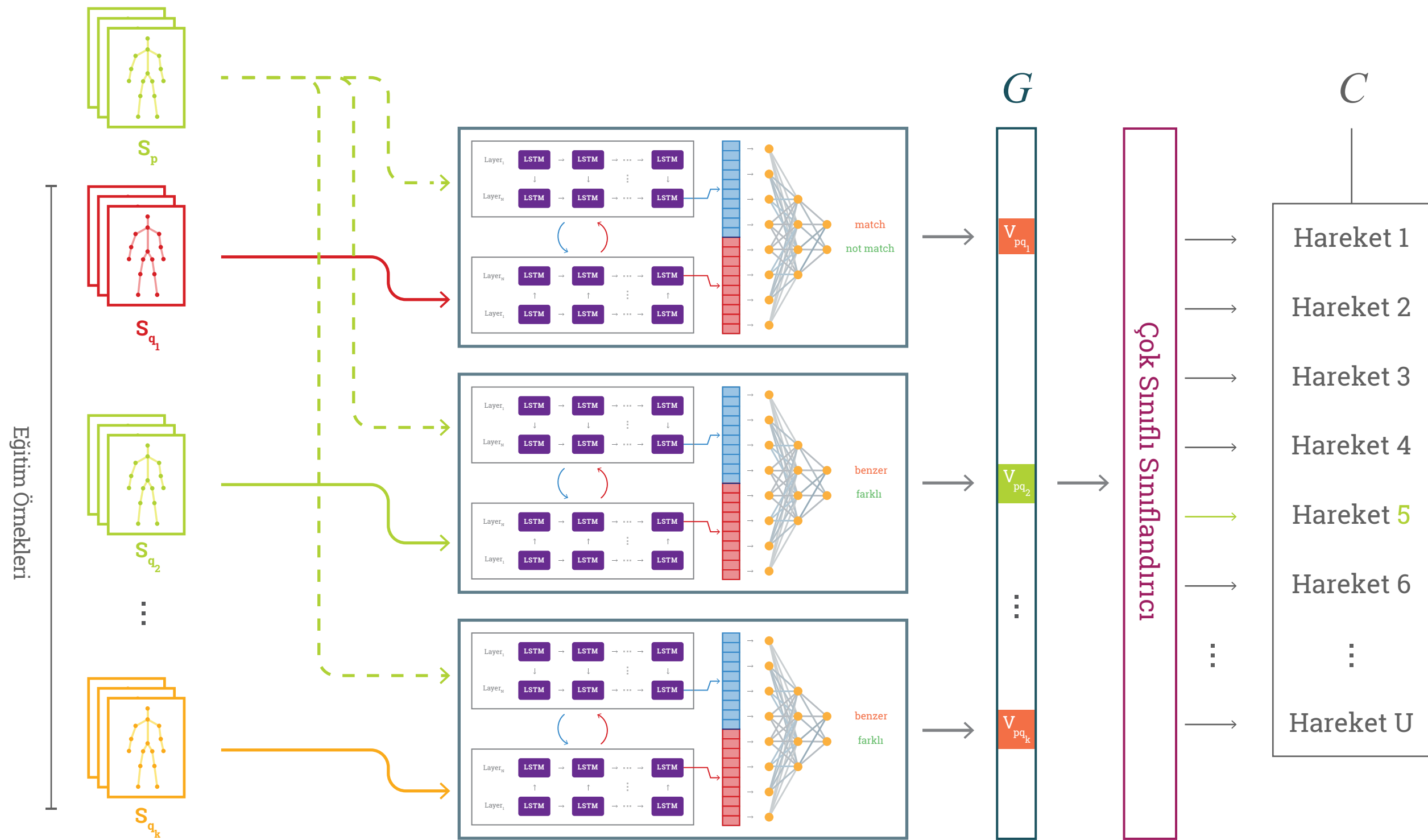


Şekil II. Geometrik Öznitelikler: ek^t , edu^t , ekf^t .



Şekil III. Kelime çantası yöntemi için kelimelerin oluşturulması.

İkiz LSTM DMÖ

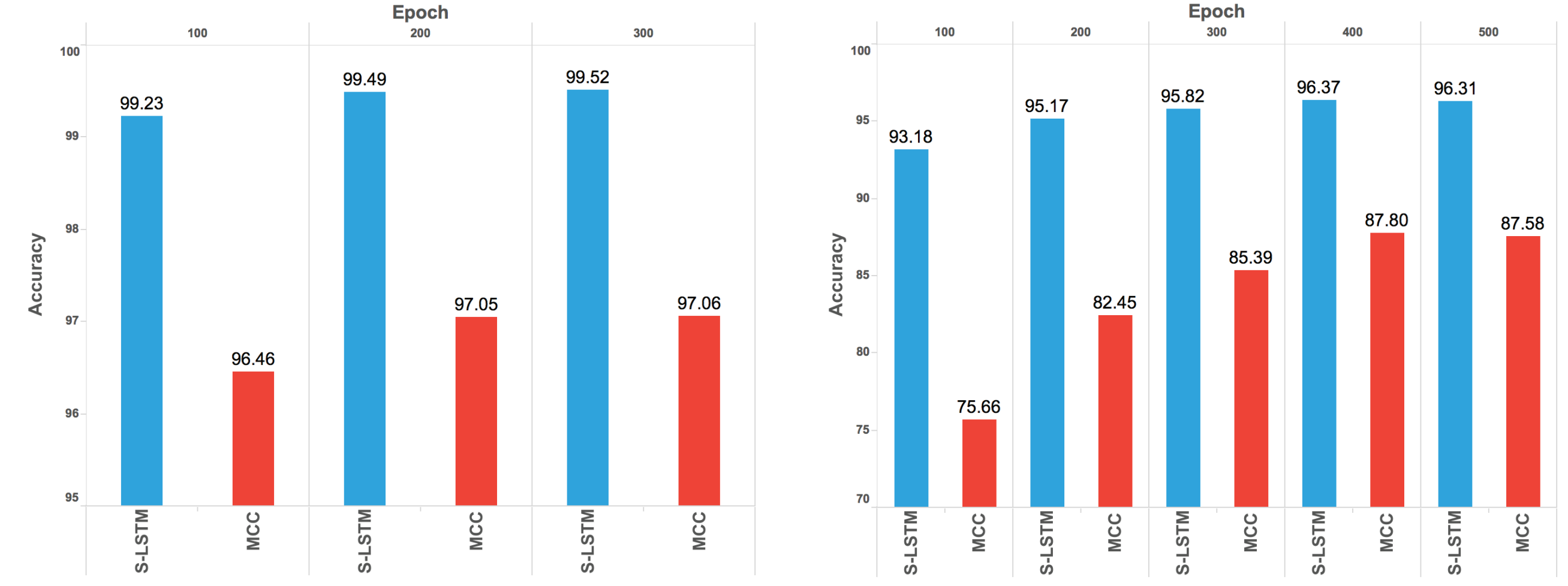


Şekil IV. İkiz LSTM DMÖ sınıflandırıcı ağı genel gösterimi.

Bu çalışmada, hareketlerin analizi ve sınıflandırması için oluşturulan İkiz LSTM ağı sunulmuştur. Derin metrik öğrenme tabanlı çalışmanın amacı hareketler arasındaki metrikleri otomatik olarak öğrenen özgün bir model oluşturmaktır. Tasarlanan model İkiz LSTM-DML, verilen iki hareket girdisi arasındaki ilişkiye ait derin metrikleri çıkarmaktadır. Hareketler ikiz ağda karşılıklı kıyaslanırken model eğitilmektedir. Hareketlerin zamansal ve uzamsal özniteliklerini modelin iç yapısı otomatik olarak çıkarmaktadır.

Sonraki bölümde paylaşılan çalışmaların sonuçlarında görülmüştür ki paylaşımlı ağlar sadece doğrulama problemleri için değil sınıflandırma problemleri için de başarılı bir şekilde çalışabilmektedir. Yapının derin metrikleri veriye uygun bir ağ ile başarılı bir şekilde çıkarılırsa, bu metrikler sınıflandırma dışında farklı uygulamalar içinde kullanılabilir. Tezin sonrasında yapılabilecek çalışmalar; bu modelin uçtan uca sınıflandırıcı olarak çalışan yapısının iyileştirilmesi, farklı veri kümelerinin bir arada kullanılması ve öğrenilmemiş bir hareketin sınıflandırılmaya çalışılması şeklinde sıralanabilir.

DeneySEL Sonuçlar



Şekil V. GTU Action 3D ve Florence Action 3D veri kümesi üzerinde İkiz LSTM DMÖ Ağı ile İkiz LSTM DMÖ sınıflandırıcı ağının başarımlarının karşılaştırılması.

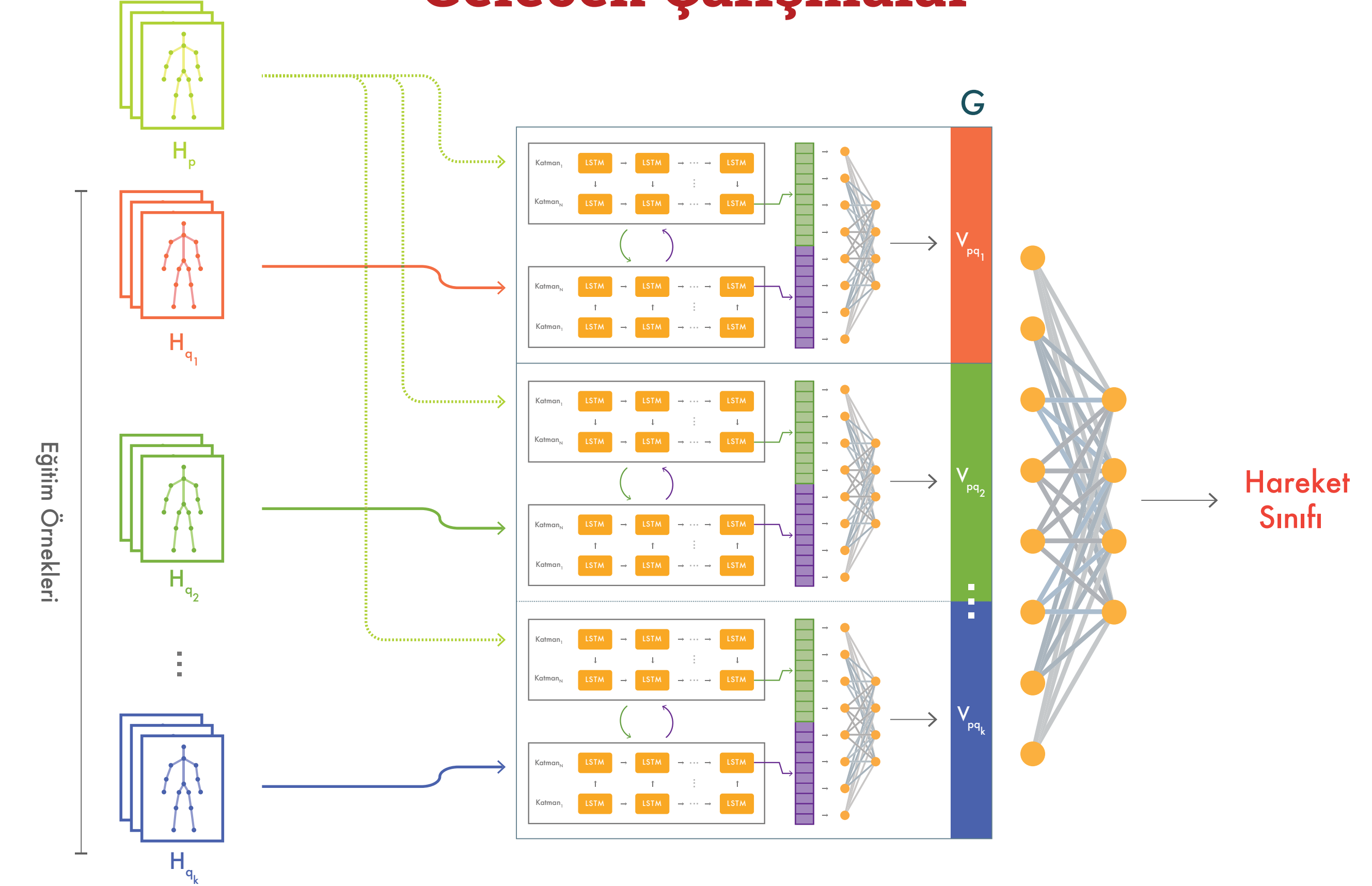
TABLO I. Florence Action 3D veri kümesi başarımları.

Metot	Başarım
Geometrik eklem çantası	88.00%
Multi-part Bag-of-Poses [2]	82.00%
Riemannian Manifold [3]	87.04%
Latent Variables [4]	89.67%
Lie Group [5]	90.88%
Feature Combinations [6]	94.39%
Softmax	61.61%
1 katmanlı LSTM	76.99%
2 katmanlı LSTM	72.32%
İkiz-LSTM DMÖ	89.51%

TABLO II. GTU Action 3D veri kümesi başarımları.

Standart Metotlar	Başarım
Geometrik eklem çantası	96.50%
SOFTMAX	75.99%
1 katmanlı LSTM	90.21%
2 katmanlı LSTM	95.47%
İkiz-LSTM DMÖ	97.06%

Gelecek Çalışmalar



Şekil VI. Uçtan Uca İkiz LSTM DMÖ sınıflandırıcı ağı genel gösterimi.

Referanslar

- [1] Skepxels: Spatio-temporal Image Representation of Human Skeleton Joints for Action Recognition - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/Skeleton-and-RGB-sample-frames-from-the-NTU-RGB-D-Human-Activity-Dataset-30-Three_fig3_321125021 [accessed 14 May, 2018]
- [2] L. Seidenari, V. Varano, S. Berretti, A. Del Bimbo, ve P. Pala, "Recognizing actions from depth cameras as weakly aligned multi-part bag-of-poses", içinde IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, 2013, ss. 479–485.
- [3] M. Devanne, H. Wannous, S. Berretti, P. Pala, M. Daoudi, ve A. Del Bimbo, "3-D Human Action Recognition by Shape Analysis of Motion Trajectories on Riemannian Manifold", IEEE Trans. Cybern., c. 45, sayı 7, ss. 1340–1352, 2015.
- [4] R. Anirudh, P. Turaga, J. Su, ve A. Srivastava, "Elastic functional coding of human actions: From vector-fields to latent variables", içinde Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2015, c. 07–12–June, ss. 3147–3155.
- [5] R. Vemulapalli, F. Arrate, ve R. Chellappa, "Human action recognition by representing 3D skeletons as points in a lie group", içinde Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2014, ss. 588–595.
- [6] D. Carbonera Luvizon, H. Tabia, ve D. Picard, "Learning features combination for human action recognition from skeleton sequences", Pattern Recognition Letters, 2016.
- [7] D. Yi, Z. Lei, ve S. Z. Li, "Deep Metric Learning for Practical Person Re-Identification", Icp, c. 11, sayı 4, ss. 1–11, 2014.

