Voice classification using Python and Keras

Cesar Osorio

PyCon 2019

I am Cesar

- Data Scientist at MoneyKey, an online lender.
- I worked on various projects related to fraud detection and risk assessment using machine learning and graph networks
- CS master’s student at Georgia Institute of Technology

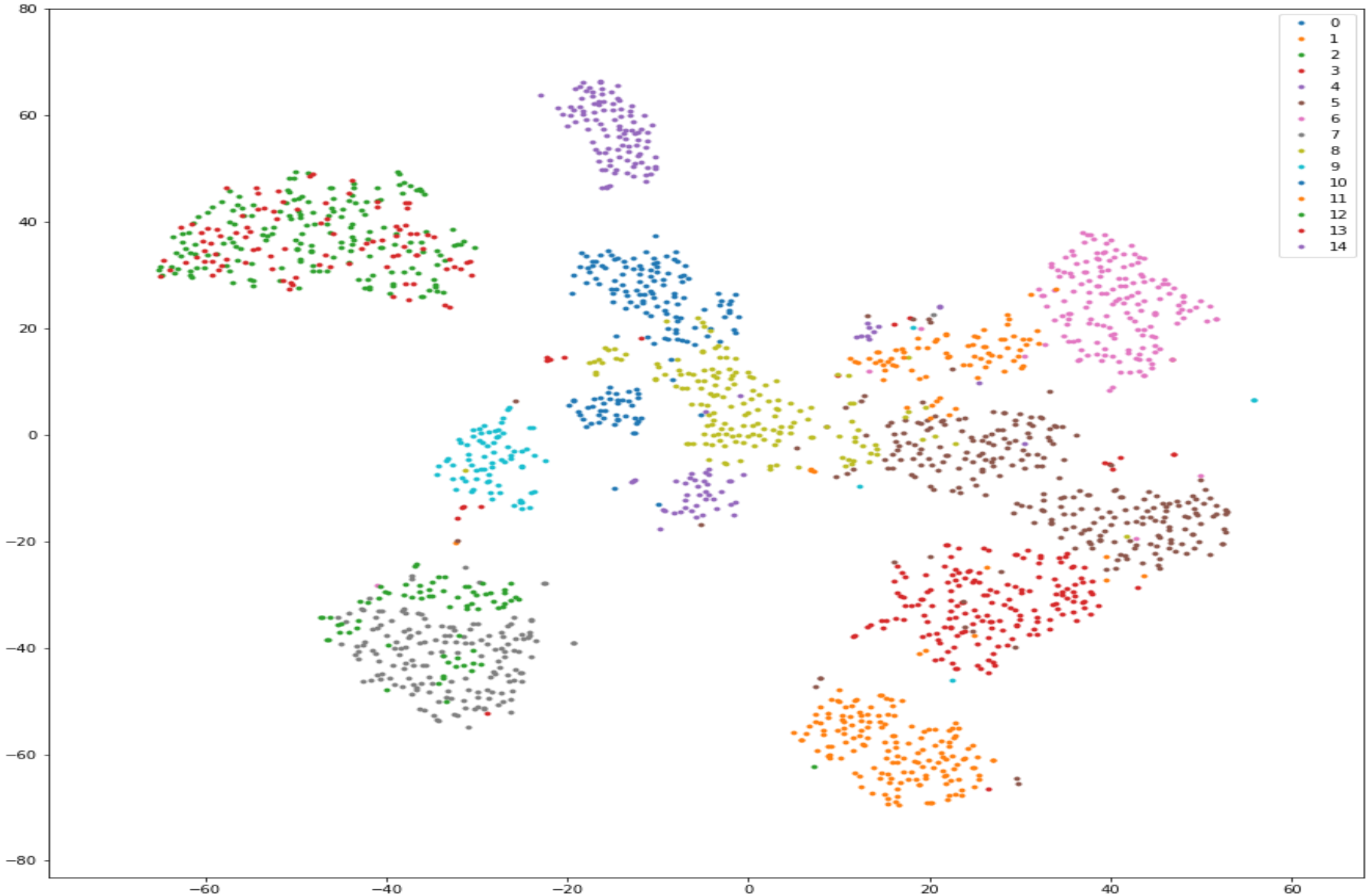
Repository: pycon2019_voice

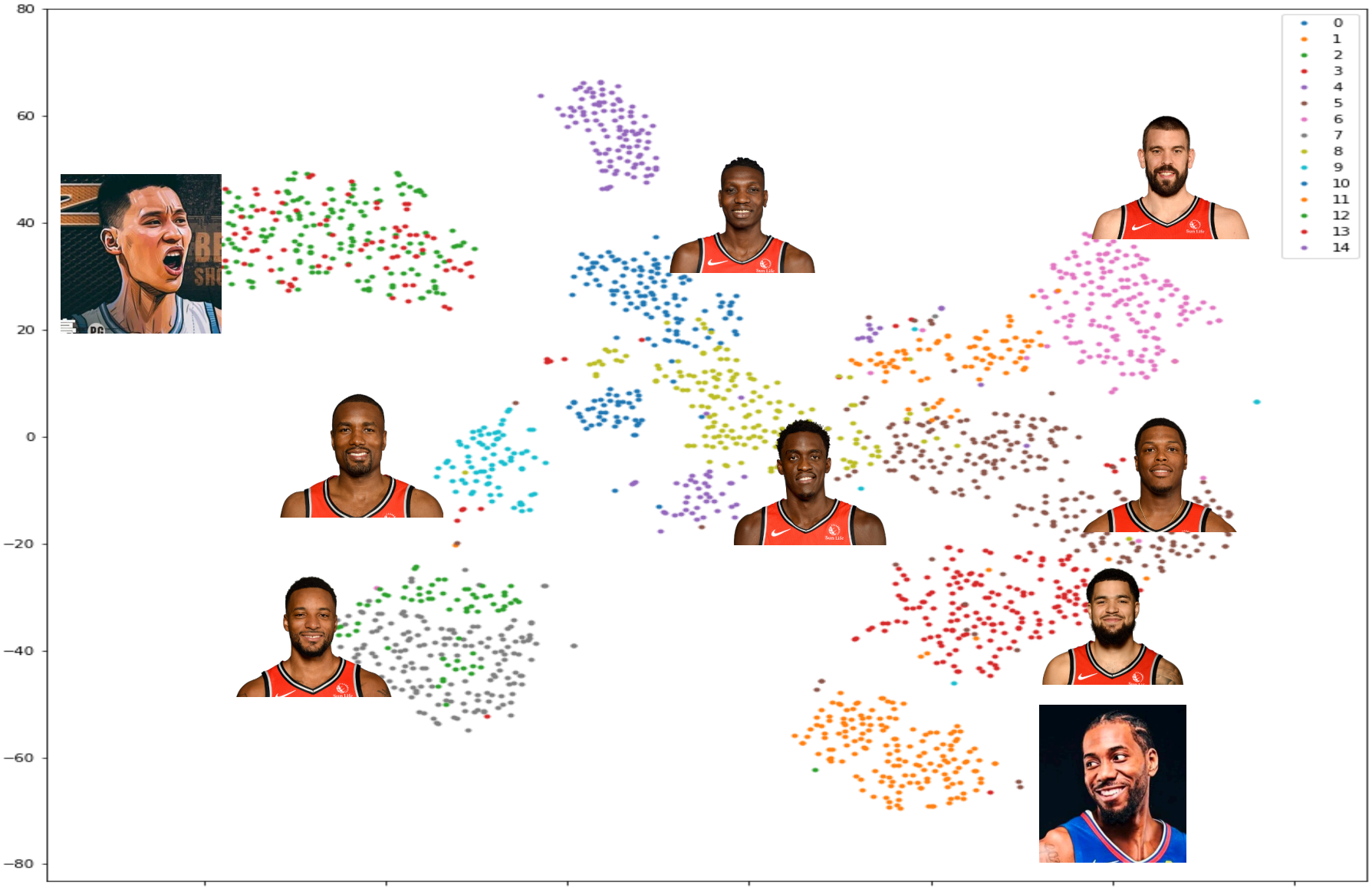
https://github.com/Cesaros

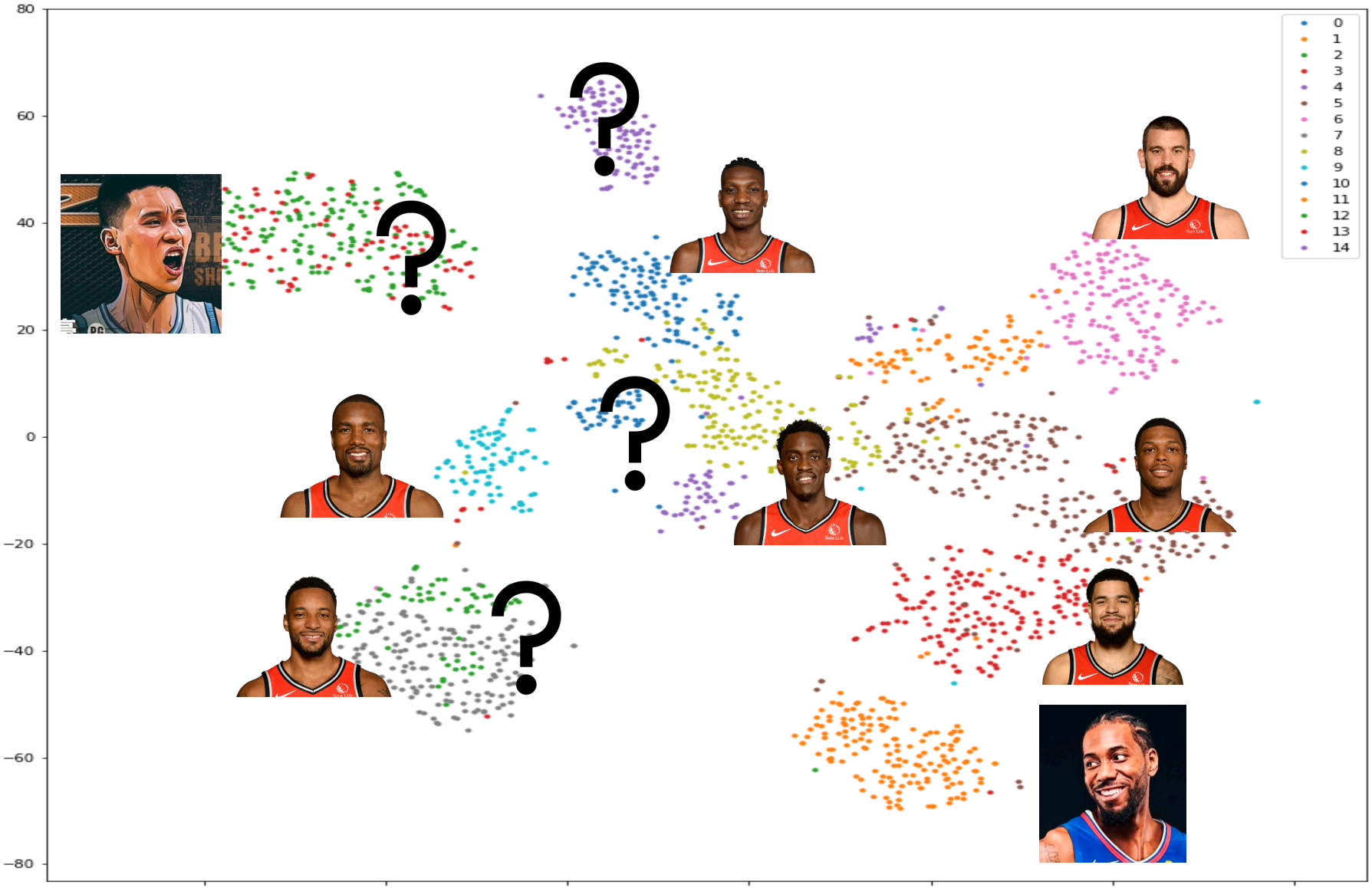
Firstname = label

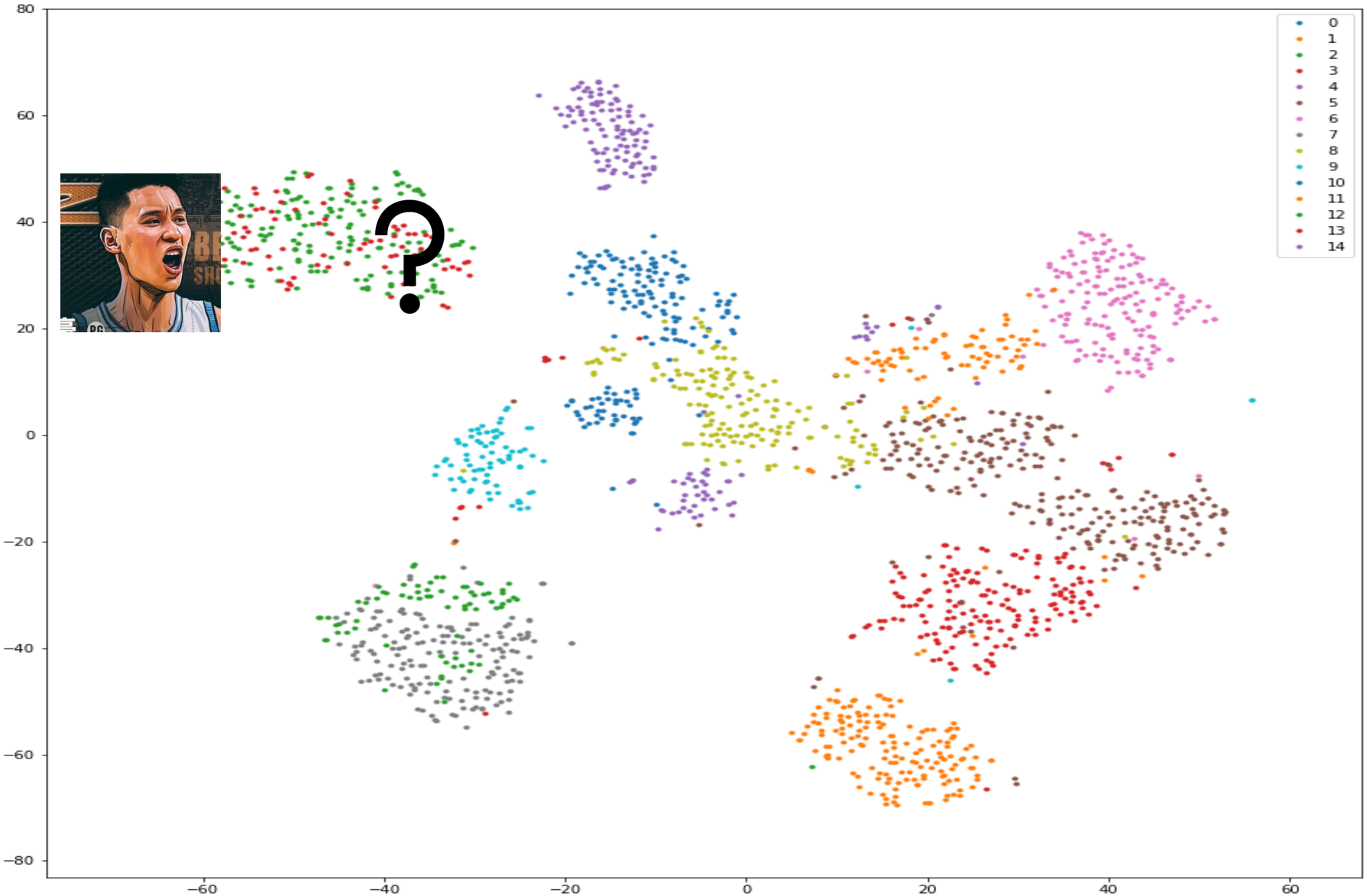
Mislabeled

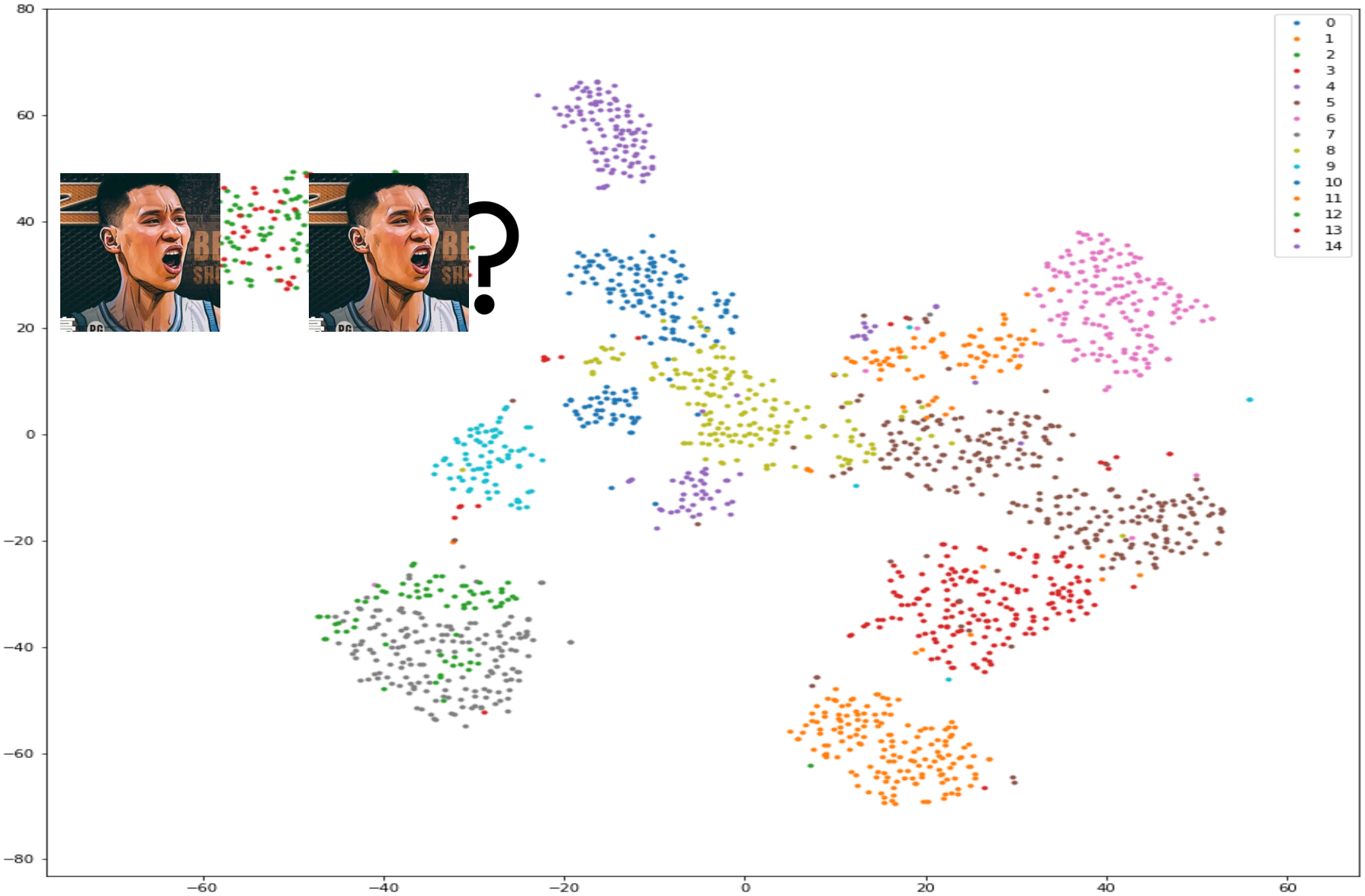
	file_name	firstname	lastname	language	raptor	length
0	kyle3_1.wav	Kyle	Lowry	english	yes	128.02
1	kyle4_1.wav	Kyle	Lowry	english	yes	134.43
2	pascal2_1.wav	Pascal	Siakam	english	yes	73.67
3	pascal3_2.wav	Pascal	Siakam	english	yes	90.88
4	marc2_1.wav	Marc	Gasol	english	yes	169.34
5	kawhi1_1.wav	Kawhi	Leonard	english	no	182.64
6	fred1_1.wav	Fred	VanVleet	english	yes	164.97
7	norman1_2.wav	Norman	Powell	english	yes	165.67
8	jeremy1_2.wav	Jeremy	Lin	english	no	180.39
9	kevinsmith1_1.wav	Kevin	Smith	english	no	53.50
10	serge1_1.wav	Serge	Ibaka	english	yes	78.87
11	chris1_1.wav	Chris	Boucher	french	yes	114.46
12	kyle1_1.wav	unknown1	voice1	english	yes	46.21
13	marc1_1.wav	unknown2	voice2	english	yes	79.95
14	norman1_3.wav	unknown3	voice3	english	yes	71.92
15	jeremy1_3.wav	unknown4	voice4	english	no	100.19
16	chris2_eng_1.wav	unknown5	voice5	english	yes	101.02

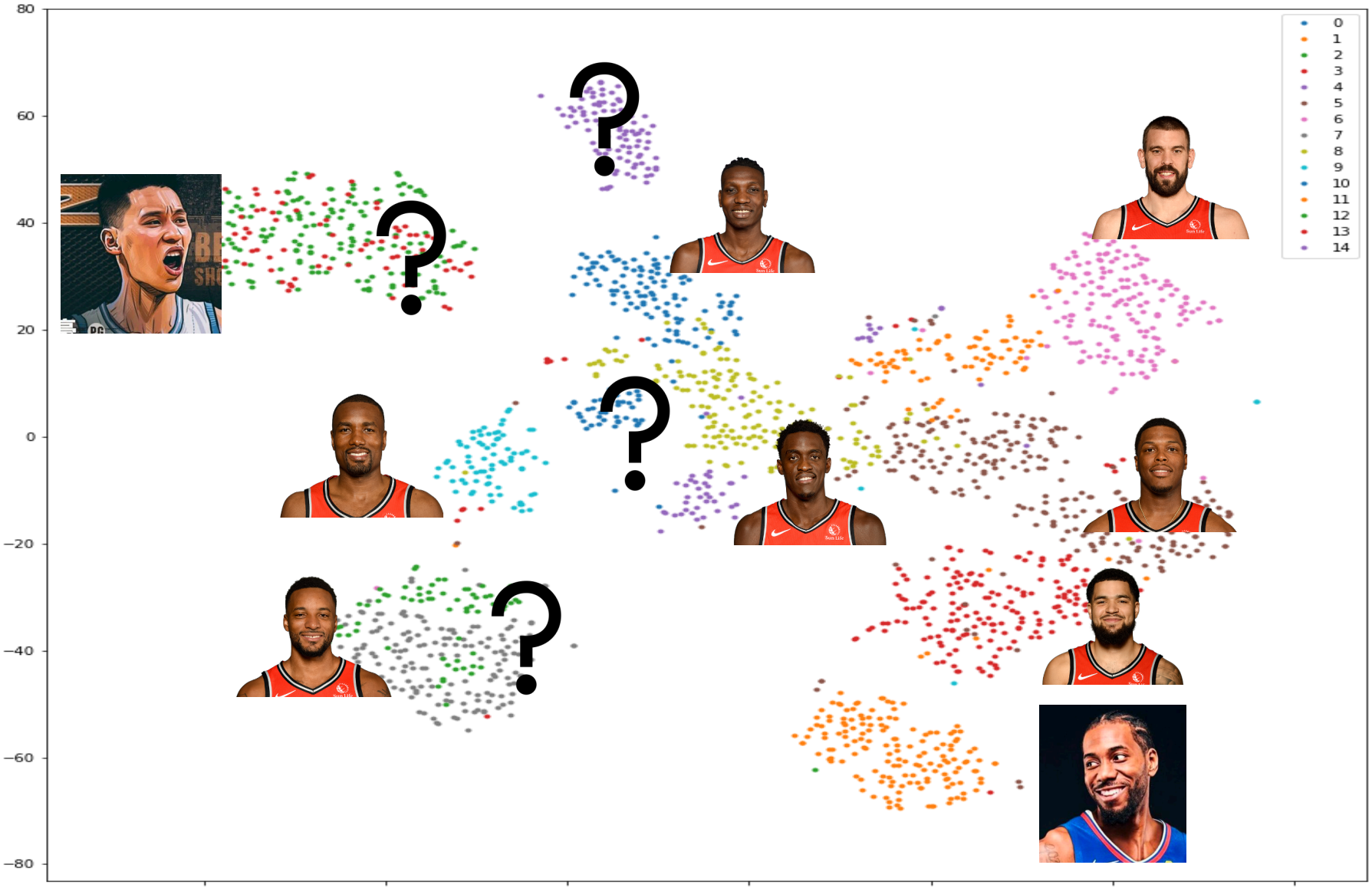

80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14

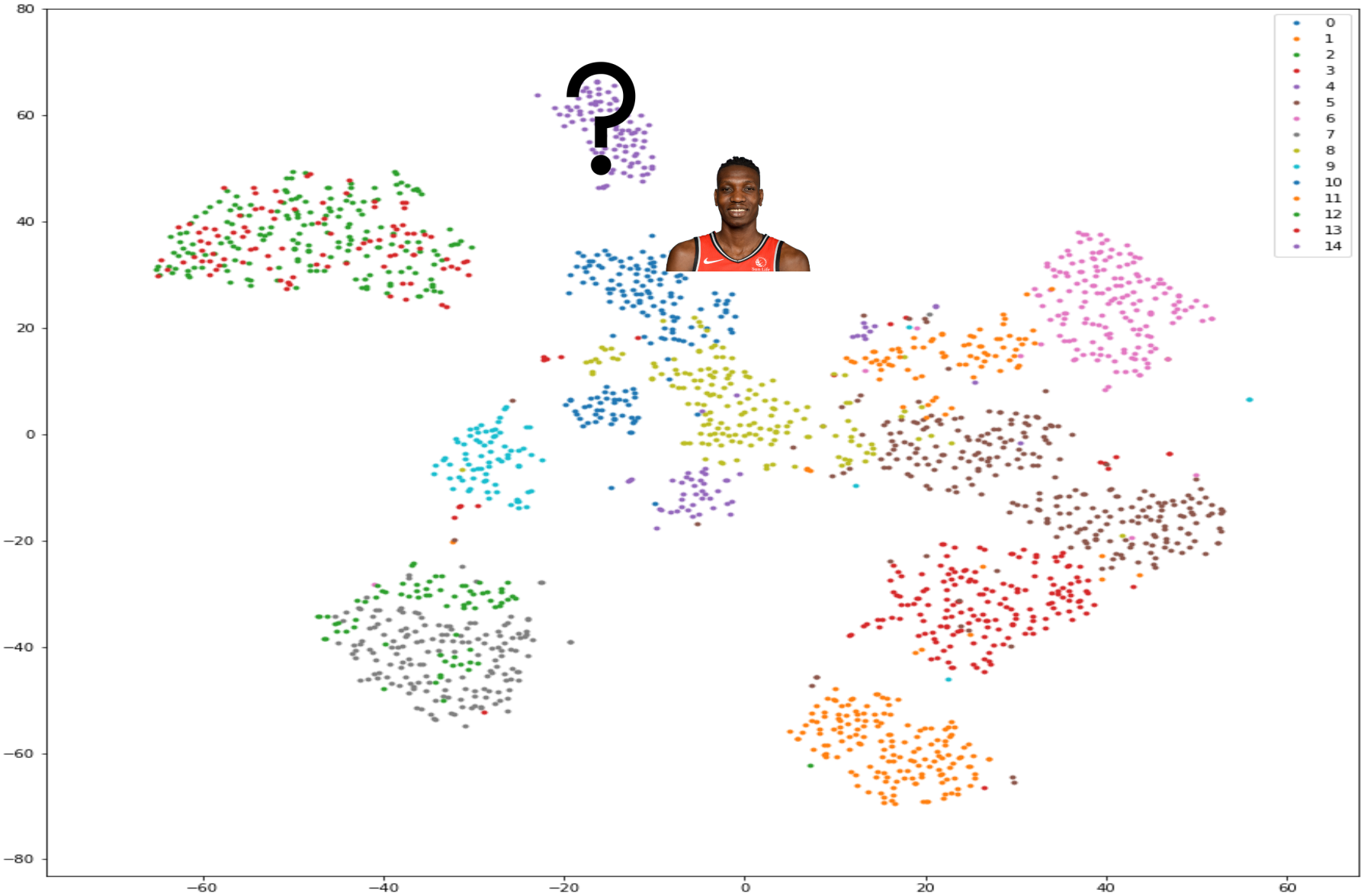

80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14

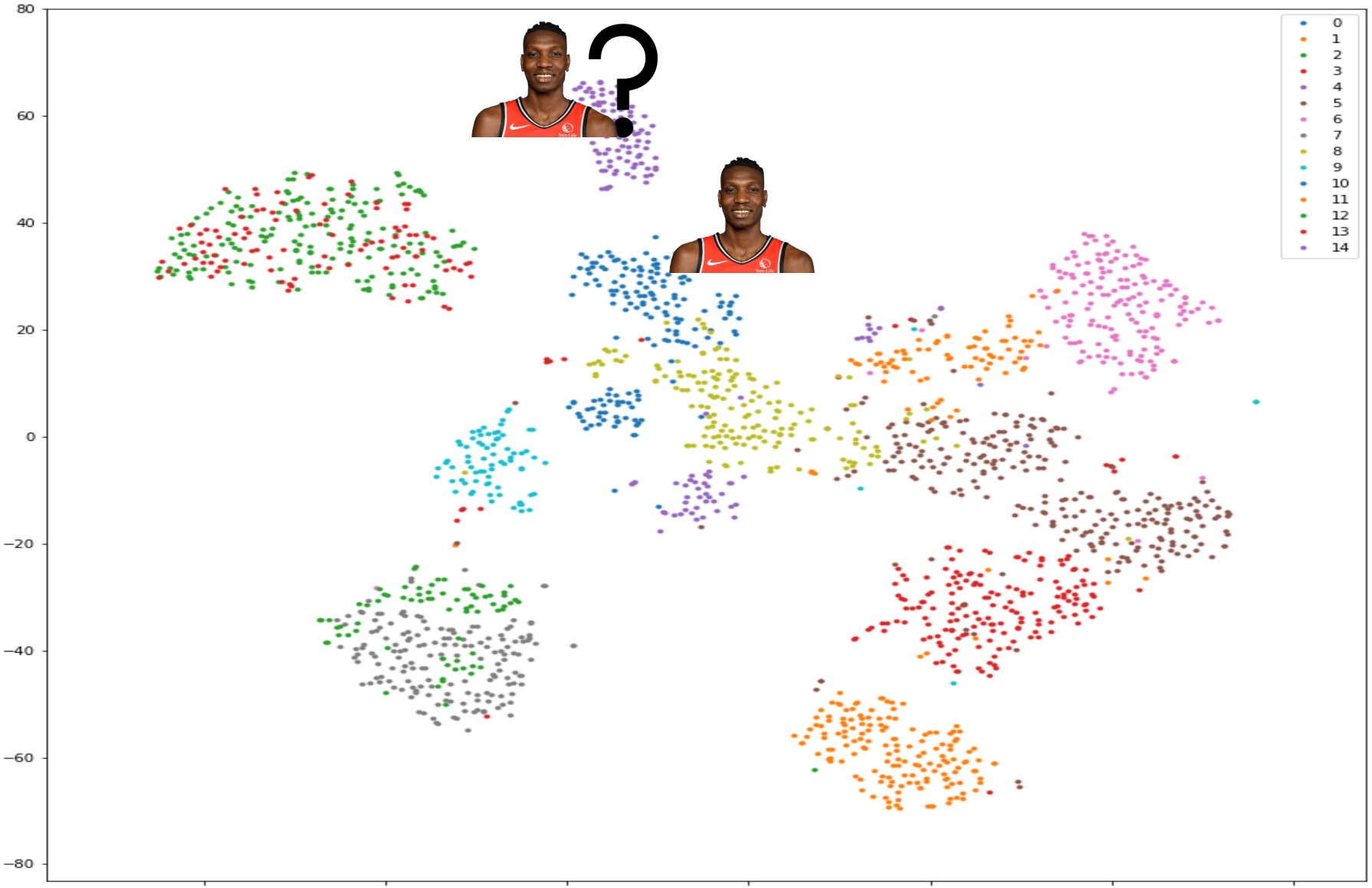

80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14

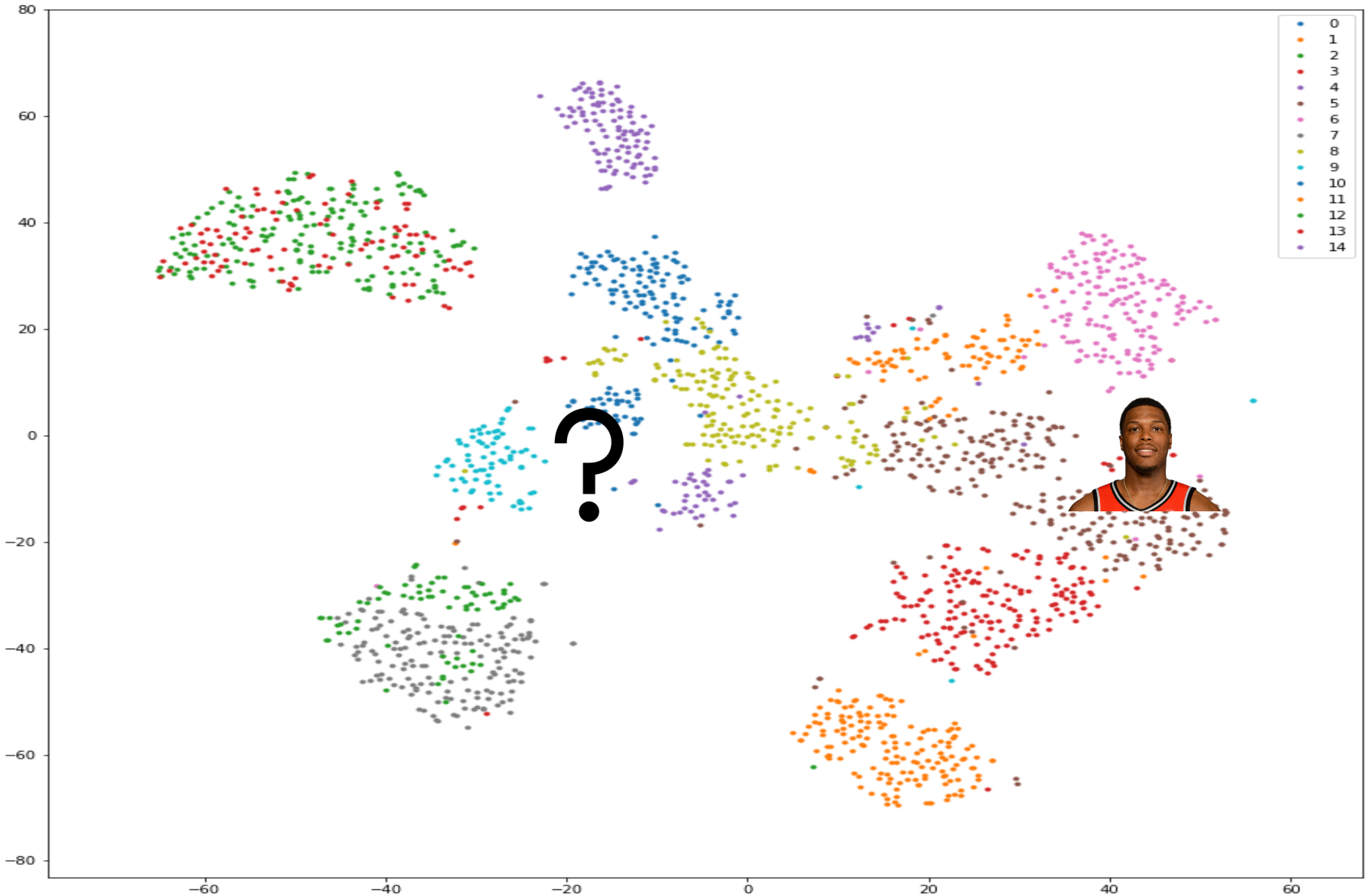

80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14

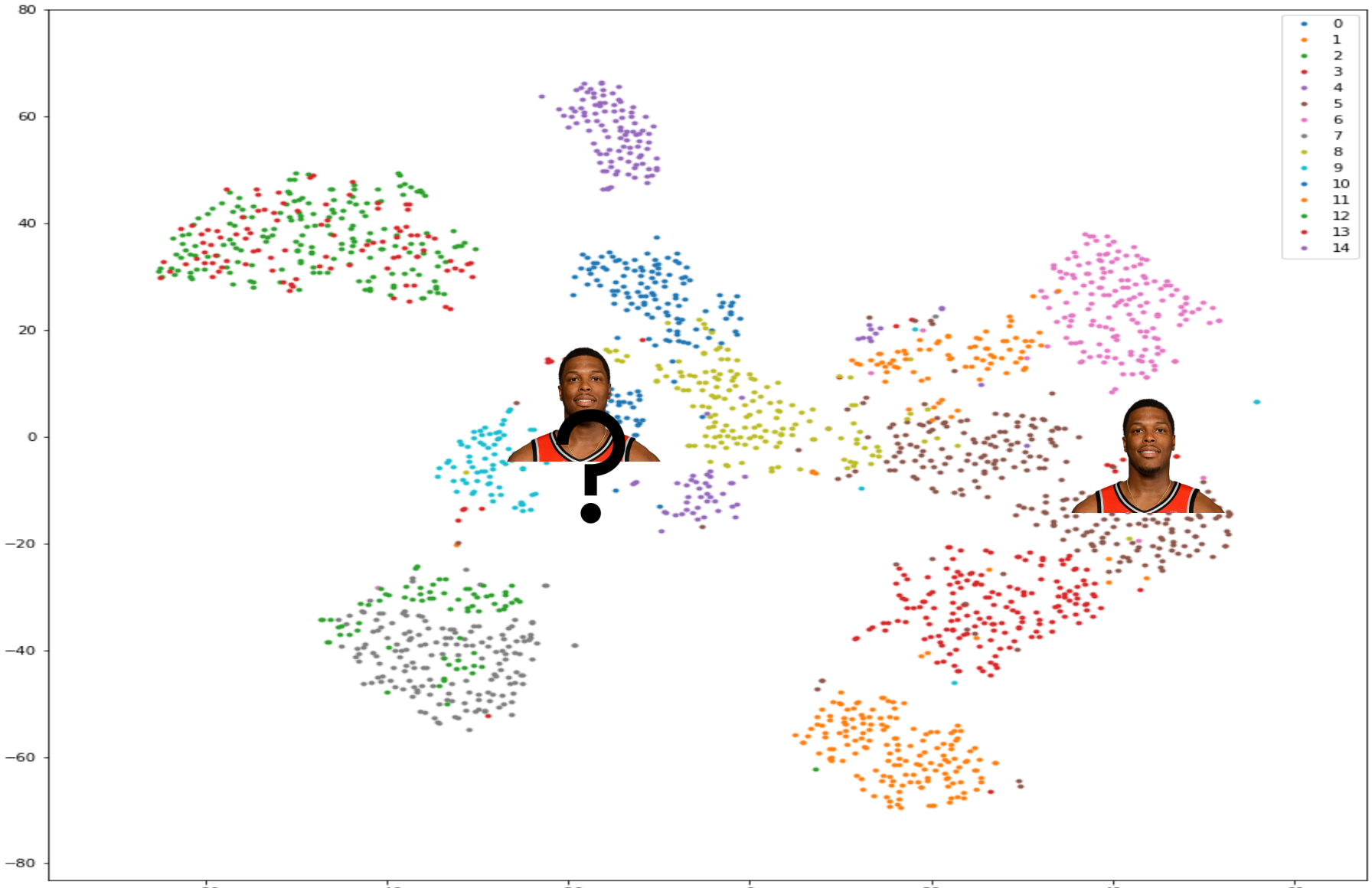

80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
?

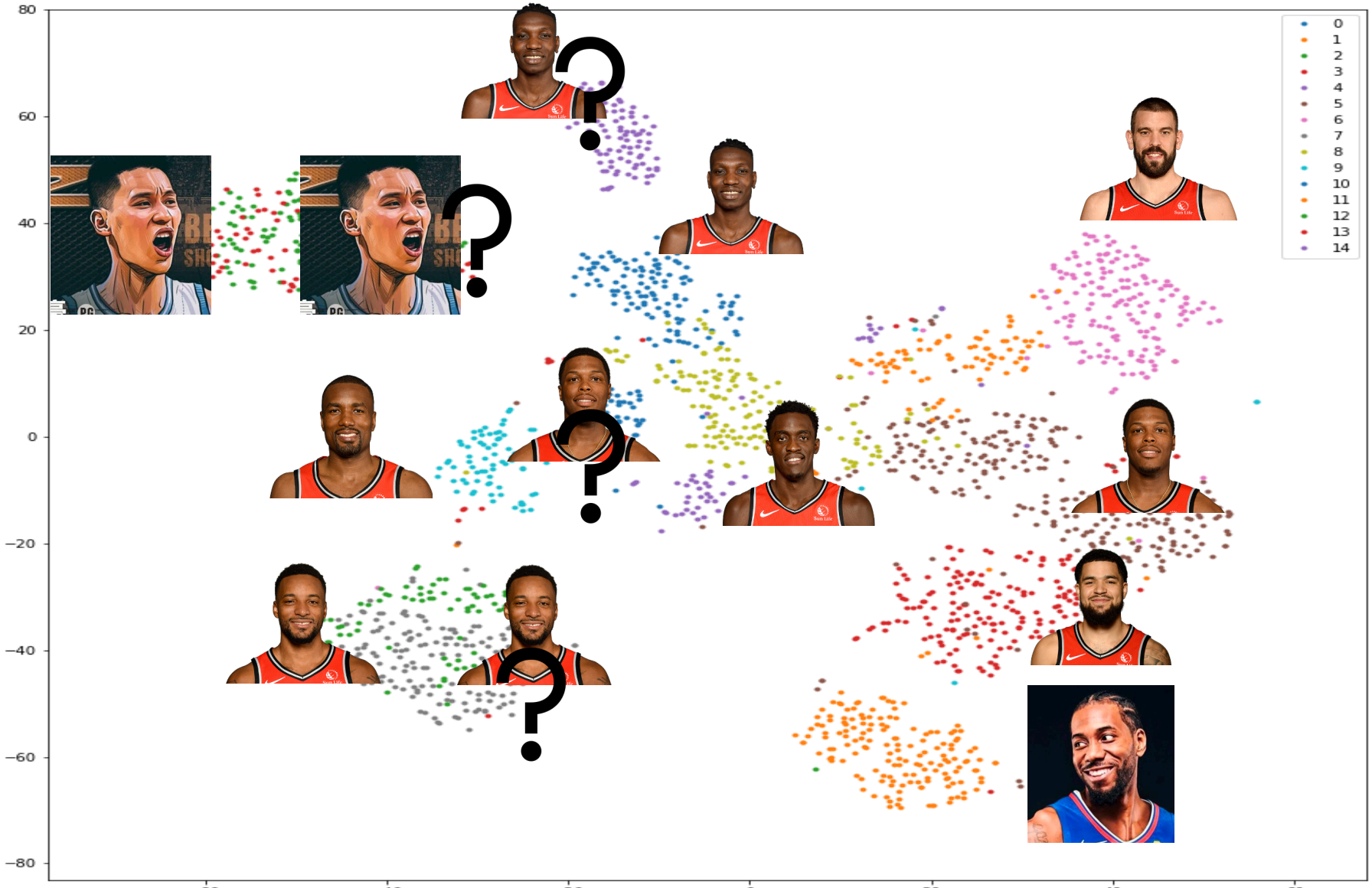

80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14


80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14


80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
?


80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14
?


80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14


80
60
40
20
0
−20
−40
−60
−80
−60
−40
−20
0
20
40
60
0
1
2
3
4
5
6
7
8
9
10
11
12
13
14

Twitter: @cesarosoriof
Github: https://github.com/Cesaros

Thanks!!