

Interprétation des résultats

Nous avons donc entraîné avec le fichier `train_resnet50.py` des réseaux de neurones sur nos différentes tâches et nous arrivons à des métriques satisfaisantes. En effet, l'accuracy, une métrique utile pour évaluer les performances d'un modèle de classification, atteint des bonnes valeurs pour chacune de nos tâches respectives.

Pour la tâche du style artistique, cette dernière contenait 27 classes et le modèle arrive à quasiment 60% d'accuracy sur les images de validation. Cela signifie que le modèle, en regardant une peinture qu'il n'a jamais vu, prédit 60% du temps le bon style artistique, parmi les 27 possibles ! Cela peut sembler peu à première vue, mais la tâche est très complexe et un modèle complètement aléatoire aurait environ 20% d'accuracy sur cette tâche (car la classe la plus représentée correspond à environ 20% des peintures), nous sommes 3 fois meilleurs !

Également, la tâche de reconnaître le genre d'un tableau (paysage, portrait, natures morte etc..) atteint 75% d'accuracy parmi ces 10 classes.

La tâche consistant à reconnaître l'artiste a été entraîné à reconnaître 23 artistes différents et nous atteignons une accuracy de plus de 60%.

Enfin, la tâche de la nationalité atteint 80% pour les 8 différentes nationalités des artistes.

Mais ce n'est pas ce qui nous intéresse.

En effet, suite à l'entraînement de ces modèles, en exécutant `feature_plots.py` nous pouvons enfin afficher le fameux feature space représentant nos données. Et l'analyser !

Nous allons prendre l'exemple de la tâche style, et nous colorions ensuite nos données selon le genre et ensuite l'artiste des tableaux, et ce pour le jeu de données d'entraînement et de test, comme on peut le voir ci-dessous :

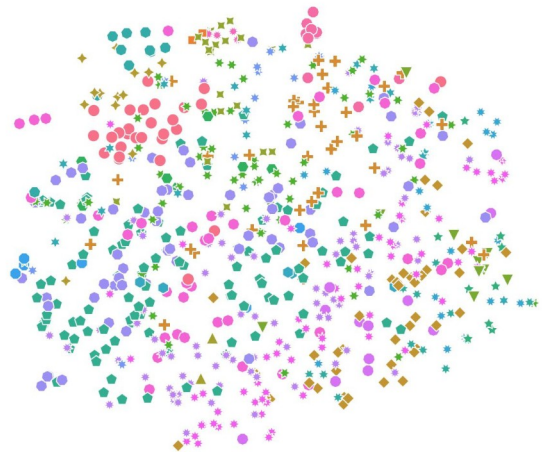
Feature space

Coloring: style

Subset: train

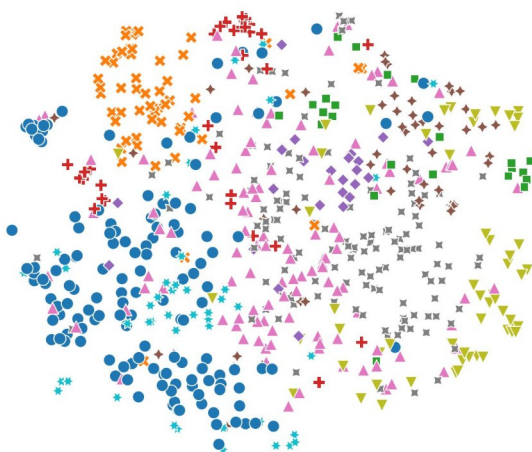


Subset: val

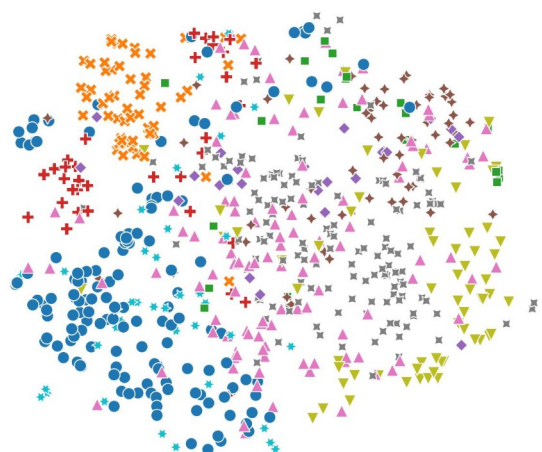


Coloring: genre

Subset: train

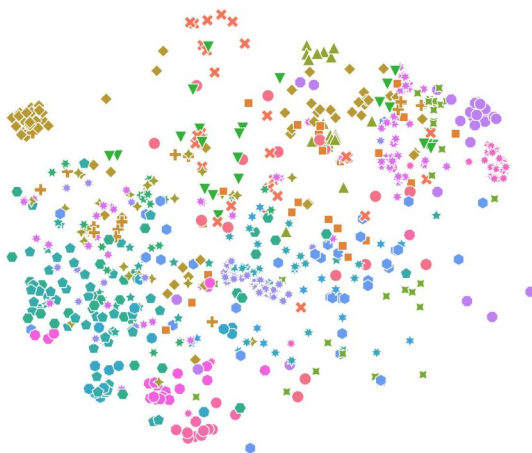


Subset: val

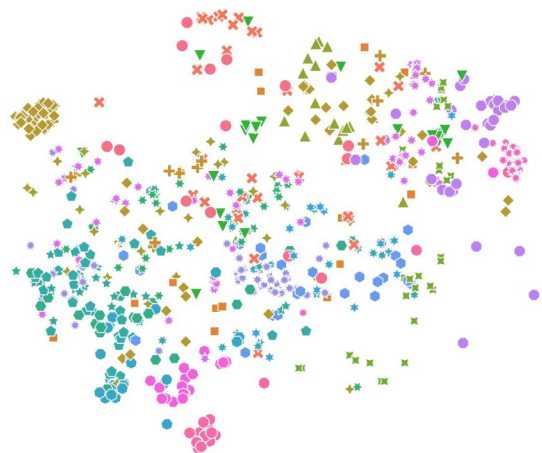


Coloring: artist

Subset: train



Subset: val



En premier lieu, si on compare les graphiques « train » et « val », on voit qu'il y a assez peu de différences. Cela signifie qu'il y a très peu d'overfit (de sur-apprentissage). On peut donc se focaliser sur n'importe lequel.

On voit en regardant le feature space coloré avec « style » (la tâche utilisée pour l'entraînement) que la tâche de classification est difficile. Cependant, on voit aussi que le modèle arrive à regrouper à certains endroits du feature space les peintures associées à des styles particuliers.

Ce qui est vraiment intéressant, c'est que le genre semble être mieux « séparé » alors qu'on n'a pas utilisé le genre pour entraîner nos modèles. On sait qu'un genre peut être associé à plusieurs styles artistiques, mais finalement, peut-être que certains styles sont plus spécifiques à certains genres ? C'est ce que laisse penser le feature space vu ici.

Enfin les artistes, eux aussi, sont assez bien séparés.