

Le formulaire en Français:

Décrivez le projet en quelques phrases. Il ne doit PAS s'agir d'une revue, mais simplement d'une description du problème abordé par le projet, de sa solution et d'un résumé des résultats obtenus.

Le but du projet est de séparer différentes sources audio.

L'équipe utilise les systèmes ECAPA-TDNN et Sepformer. ECAPA-TDNN est un modèle basé sur l'attention qui permet d'extraire des embeddings. Sepformer permet de séparer des sources audio. Ecapa-TDNN pour l'extraction des embeddings, prend des segments audio et extraire le context local et globales.

Elle a fait une revue de littérature et identifié le conditionnement FiLM.

L'équipe propose d'extraire des embeddings à l'aide du modèle Ecapa-TDNN pour les injecter dans sepformer.

Le projet utilise le dataset: LibriSpeech, 470h, 1252 locuteurs.

Citez et justifiez au moins deux aspects forts du projet.

Utilise le calcul de similarité sur les embeddings des voix.

Basé sur la littérature existante

Utilise une méthode intéressante d'augmentation de données en injectant des embeddings dans les données présentées à Sepformer

Énumérez et justifiez au moins deux aspects manquants du projet.

Pas de présentation des résultats. Il aurait été intéressant de voir des résultats; même si ceux-ci sont partiels. Ça aurait permis d'entrevoir le potentiel du projet.

Personnalisation du modèle original manquante. Le développement d'un modèle original aurait permis de démontrer un avancement sur l'état de l'art.

Énumérez au moins trois questions à l'équipe du projet.

Comment avez-vous trouvé votre modèle? L'avez vous codé où est-il accessible sur le net open source?

Pensez vous montrer les résultats sur le calcul de similarité sur le même locuteur?

Avez-vous évalué la possibilité de développer votre propre modèle?