



Université Sultan Moulay Slimane École Nationale des Sciences Appliquées - Khouribga

Filière : Informatique et Ingénierie des Données $Ann\acute{e}e~universitaire~2024-2025$

RAPPORT DE PROJET

détection des émotion l'hors d'une vidéo-conférence

Réalisé par :

- Mohamed Rida Lajghal
- Abdellah Doulfaqar
- Moad Chergui

Encadré par :

Professeur Nidal LAMGHARI

TABLE DES MATIÈRES

1	Intr	roduction	6
	1.1	Contexte et Problématique	6
		1.1.1 L'Évolution de l'Enseignement à Distance	6
		1.1.2 Défis Liés à l'Engagement des Étudiants dans l'EAD	6
		1.1.3 Lacunes des Solutions Actuelles de Visioconférence	6
		1.1.4 Opportunité de la Détection des Émotions	7
	1.2	Objectifs du Projet	7
		1.2.1 Objectifs Généraux	7
		1.2.2 Objectifs Spécifiques	7
2	Ana	alyse et Conception	9
	2.1	Analyse des Besoins	9
		2.1.1 Besoins Fonctionnels	9
		2.1.2 Besoins Non Fonctionnels	10
	2.2	Conception de l'Architecture	11
		2.2.1 Architecture Générale du Système	11
		2.2.2 Diagrammes UML	12
3	Imp	olémentation	17
	3.1	Introduction	17
	3.2	Choix Technologiques	17
	3.3	Architecture Générale du Système	18
		3.3.1 Composants du Système	18
	3.4	Fonctionnalités Implémentées	18
	3.5	Conclusion	19

4	Rés	ultats		20
	4.1	Conne	exion	20
		4.1.1	Authentification	20
		4.1.2	Page d'accueil	21
	4.2	Gestio	on des Réunions	22
		4.2.1	Création d'une Réunion	22
		4.2.2	Rejoindre une Réunion	23
		4.2.3	Déroulement d'une Réunion	23
	4.3	Histor	ique des Réunions	24
		4.3.1	Réunions créées	24
		4.3.2	Réunions rejointes	25
5	Con	clusio	n et Perspectives	27

TABLE DES FIGURES

2.1	Diagramme de Cas d'Utilisation	14
2.2	Diagramme de Séquence (connexion et authentification)	15
2.3	Diagramme de Séquence (déroulement de la réunion)	15
2.4	Diagrammes de Déploiement	16
3.1	Technologies utilisées dans le projet	17
4.1	Interface d'authentification Google via Firebase	21
4.2	Page d'accueil de l'application	21
4.3	Page de création de réunion	22
4.4	Création d'une réunion individuelle	22
4.5	Création d'une visioconférence	23
4.6	Interface pour rejoindre une réunion	23
4.7	Déroulement d'une réunion en cours	24
4.8	Liste des réunions créées	24
4.9	Modification d'une réunion	25
4.10	Consultation des statistiques d'une réunion	25
4.11	Réunions rejointes	26

LISTE DES TABLEAUX

CHAPITRE 1

INTRODUCTION

1.1 Contexte et Problématique

1.1.1 L'Évolution de l'Enseignement à Distance

L'avènement des technologies numériques a profondément transformé le paysage éducatif, rendant l'enseignement à distance (EAD) de plus en plus prévalent. La crise sanitaire récente a accéléré cette transition, poussant de nombreuses institutions à adopter massivement les plateformes de visioconférence pour assurer la continuité pédagogique. Si ces outils offrent une flexibilité et une accessibilité inédites, ils introduisent également de nouveaux défis.

1.1.2 Défis Liés à l'Engagement des Étudiants dans l'EAD

Malgré les avantages de l'EAD, maintenir l'engagement et la motivation des étudiants à distance est un défi majeur. L'absence d'interaction physique directe et de signaux non verbaux tangibles rend difficile pour les enseignants de percevoir l'état émotionnel de leurs élèves. Un étudiant peut sembler attentif mais être en réalité distrait, frustré ou en difficulté de compréhension, sans que l'enseignant ne puisse le détecter aisément. Cette lacune peut nuire à l'efficacité de l'apprentissage et à la qualité de l'enseignement.

1.1.3 Lacunes des Solutions Actuelles de Visioconférence

Les plateformes de visioconférence existantes sont principalement axées sur la communication audio-visuelle et le partage de contenu. Elles intègrent rarement des fonctionnalités avancées permettant d'analyser l'état émotionnel des participants. Cette absence de retour d'information émotionnel en temps réel limite la capacité des enseignants à adapter leur pédagogie, à identifier les étudiants en décrochage ou à évaluer l'impact de leurs explications. Il devient donc impératif de développer des outils complémentaires pour combler cette lacune.

1.1.4 Opportunité de la Détection des Émotions

La détection automatique des émotions, grâce aux avancées en intelligence artificielle et en vision par ordinateur, offre une opportunité prometteuse pour enrichir l'expérience de l'EAD. En analysant les expressions faciales, les mouvements de tête et d'autres indicateurs visuels, il est possible d'inférer l'état émotionnel des étudiants. L'intégration d'une telle fonctionnalité dans une application de visioconférence pourrait fournir aux enseignants des informations précieuses pour mieux comprendre leurs élèves, personnaliser leur approche et améliorer l'environnement d'apprentissage virtuel.

1.2 Objectifs du Projet

1.2.1 Objectifs Généraux

L'objectif général de ce projet est de concevoir et de développer une application de visioconférence innovante intégrant un module de détection des émotions des étudiants en temps réel. Cette intégration vise à fournir aux enseignants des outils améliorés pour comprendre l'engagement et l'état émotionnel de leurs élèves lors des sessions d'enseignement à distance. En identifiant les émotions telles que la confusion, l'ennui ou l'intérêt, l'application cherchera à permettre une pédagogie plus adaptative et personnalisée, favorisant ainsi une meilleure expérience d'apprentissage.

1.2.2 Objectifs Spécifiques

Pour atteindre notre objectif général, nous nous sommes fixés les objectifs spécifiques suivants :

- **Développer un module de visioconférence robuste et fonctionnel** : Ce module devra assurer une communication audio et vidéo stable, ainsi que les fonctionnalités de base attendues d'une plateforme d'enseignement en ligne (partage d'écran, chat, etc.).
- Intégrer un système de détection des émotions faciales en temps réel : Le système devra être capable d'analyser les expressions faciales des étudiants et d'inférer leurs émotions (joie, tristesse, surprise, colère, dégoût, peur, neutre, confusion, ennui, etc.) avec une précision acceptable.
- Fournir aux enseignants un tableau de bord visuel des émotions : Ce tableau de bord devra présenter de manière agrégée et intuitive les émotions dominantes au

- sein du groupe d'étudiants, et potentiellement permettre de visualiser les émotions individuelles de manière anonymisée ou non, selon la configuration.
- Assurer la performance et la réactivité du système : La détection des émotions et la visualisation des données doivent être suffisamment rapides pour être utiles en temps réel et ne pas perturber l'expérience de visioconférence.

CHAPITRE 2

ANALYSE ET CONCEPTION

2.1 Analyse des Besoins

L'analyse des besoins est une étape cruciale qui permet de définir précisément ce que l'application doit faire et comment elle doit se comporter pour répondre aux attentes des utilisateurs finaux (enseignants et étudiants). Cette section détaille les besoins fonctionnels et non fonctionnels identifiés pour notre application de visioconférence avec détection d'émotions.

2.1.1 Besoins Fonctionnels

Les besoins fonctionnels décrivent les fonctionnalités spécifiques que le système doit offrir. Pour notre application, ils incluent :

Gestion des Sessions de Visioconférence

- Création et Join de Sessions : Les enseignants doivent pouvoir créer des sessions de cours en ligne, générer des liens ou des codes d'invitation, et les étudiants doivent pouvoir rejoindre ces sessions facilement.
- Communication Audio et Vidéo : Assurer une transmission fluide de l'audio et de la vidéo entre tous les participants, avec des options de coupure de microphone et de caméra.
- Partage d'Écran : Permettre aux enseignants de partager leur écran pour des présentations, des démonstrations ou des documents.

- Chat Intégré : Offrir une fonctionnalité de chat textuel pour les questions, les commentaires et les interactions écrites pendant la session.
- **Gestion des Participants** : L'enseignant doit pouvoir voir la liste des participants, les admettre ou les expulser.

Détection et Analyse des Émotions

- Capture Vidéo en Temps Réel : La caméra de l'étudiant doit être utilisée pour capturer des flux vidéo destinés à la détection des émotions.
- **Reconnaissance Faciale** : Le système doit être capable de détecter les visages des étudiants dans le flux vidéo.
- **Détection des Émotions Primaires**: Le modèle doit pouvoir identifier un ensemble d'émotions de base (joie, tristesse, colère, surprise, peur, dégoût, neutre).
- **Détection d'Émotions Contextuelles** : Si possible, le système devrait également chercher à identifier des émotions plus spécifiques au contexte pédagogique, comme la confusion, l'ennui, la curiosité ou la concentration.
- Agrégation des Données Émotionnelles : Les émotions détectées doivent être agrégées au niveau individuel et au niveau du groupe d'étudiants.
- Visualisation pour l'Enseignant : Un tableau de bord visuel doit présenter les émotions détectées de manière claire et compréhensible pour l'enseignant (par exemple, des graphiques en temps réel, des indicateurs d'état).

Sécurité et Confidentialité

- Authentification et Autorisation : Les utilisateurs doivent s'authentifier pour accéder aux sessions et leurs rôles doivent être gérés (enseignant, étudiant).
- Confidentialité des Données Émotionnelles : Assurer que les données émotionnelles des étudiants sont traitées de manière éthique et confidentielle, avec des options de consentement clair.

2.1.2 Besoins Non Fonctionnels

Les besoins non fonctionnels décrivent les critères de qualité du système, sa performance, sa sécurité, etc.

Performance

- Faible Latence : La transmission vidéo/audio et la détection d'émotions doivent se faire avec une latence minimale pour assurer une expérience en temps réel.
- **Scalabilité**: L'application doit pouvoir supporter un nombre croissant de participants simultanément sans dégradation significative des performances.

— **Temps de Réponse** : Le système doit être réactif aux interactions de l'utilisateur (par exemple, activation/désactivation de la caméra, partage d'écran).

Fiabilité et Robustesse

- **Stabilité** : L'application doit être stable et ne pas crasher ou rencontrer de bugs fréquents.
- Gestion des Erreurs : Le système doit gérer gracieusement les erreurs (par exemple, perte de connexion internet, échec de la détection faciale).
- **Précision de la Détection** : Le module de détection d'émotions doit présenter une précision suffisante pour être utile et fiable.

Facilité d'Utilisation (Usability)

- **Interface Intuitive**: L'interface utilisateur doit être claire, simple à comprendre et facile à naviguer pour les enseignants et les étudiants.
- **Accessibilité** : L'application doit être conçue en tenant compte des principes d'accessibilité (si applicable).

Sécurité

- **Protection des Données** : Les données personnelles et émotionnelles des utilisateurs doivent être protégées contre les accès non autorisés et les fuites.
- Cryptage des Communications : Les flux vidéo, audio et de données doivent être cryptés pour garantir la confidentialité.

Compatibilité

- Compatibilité Navigateur/OS: L'application doit fonctionner sur les principaux navigateurs web et systèmes d'exploitation (Windows, macOS, Linux).
- Compatibilité Matériel : L'application doit être compatible avec une variété de webcams et de configurations matérielles standard.

2.2 Conception de l'Architecture

2.2.1 Architecture Générale du Système

Le système est une application web développée en JavaScript, utilisant **React** et la bibliothèque d'interface **Elastic UI**. Il permet à des utilisateurs authentifiés de participer à des réunions vidéo, avec une analyse émotionnelle automatisée en temps réel.

Utilisateurs Deux types d'utilisateurs interagissent avec le système :

- **Participant** : peut rejoindre une réunion, activer sa caméra, partager son écran et discuter via le chat.
- **Modérateur** : hérite des fonctionnalités du participant et peut en plus lancer une réunion et gérer les participants.

Composants Principaux du Système Le système repose sur l'intégration de plusieurs technologies réparties entre le frontend, le backend, et des services tiers :

- Frontend (React + Elastic UI) :

- Interface utilisateur dynamique.
- Authentification via Firebase Authentication (FireAuth).
- Intégration avec **Zegocloud** pour le streaming vidéo en temps réel.
- Capture périodique de frames vidéo (toutes les 3 secondes) pour l'analyse émotionnelle.
- Stockage des prédictions d'émotions dans **Firestore**.

— Backend Flask (Python):

- Serveur léger développé en Flask.
- Reçoit les frames vidéo depuis le frontend.
- Analyse les émotions via un modèle d'apprentissage automatique.
- Retourne les prédictions au frontend.

— Zegocloud :

— Fournit les services de visioconférence (caméra, micro, écran partagé).

Communication entre composants

- Le frontend envoie une frame toutes les 3 secondes au serveur Flask via une requête HTTP POST.
- Le serveur Flask retourne les prédictions (ex : joie, surprise, colère).
- Ces prédictions sont ensuite sauvegardées dans Firestore pour consultation ultérieure par le modérateur.

L'analyse fonctionnelle et dynamique du système est détaillée à travers les diagrammes UML présentés dans les sections suivantes.

2.2.2 Diagrammes UML

Diagramme des Cas d'Utilisation

Le système comporte trois acteurs :

- Participant : acteur principal du système.
- **Modérateur** : acteur principale hérité du Participant.
- Modèle de Détection d'Émotions : acteur secondaire.

Cas d'utilisation:

1. S'authentifier

Acteur: Participant

L'utilisateur se connecte pour accéder aux fonctionnalités du système.

2. Consulter les statistiques émotionnelles des réunions précédentes

Acteur: Participant

Précondition : Doit être authentifié

Le participant peut consulter les rapports sur les émotions détectées lors des réunions passées.

3. Rejoindre une réunion

Acteur: Participant

Précondition : Doit être authentifié

Le participant rejoint une réunion via un lien ou après sélection depuis la base de données des étudiants.

4. Lancer une réunion

Acteur : Participant ou Modérateur

Précondition : Doit être authentifié

Permet de démarrer une nouvelle réunion.

5. Partager la vidéo de la caméra

Acteur: Participant

Précondition : Pendant une réunion

Le participant peut activer le flux de sa caméra.

6. Partager l'écran

Acteur: Participant

Précondition : Pendant une réunion

Le participant peut partager son écran.

7. Envoyer des messages dans le chat

Acteur: Participant

Précondition : Pendant une réunion

Le participant échange des messages textuels via le chat intégré.

8. Gérer les participants

Acteur : Modérateur

Précondition : Pendant une réunion

Le modérateur ajoute ou retire des participants et gère leurs permissions.

9. Consulter les statistiques émotionnelles en temps réel

Acteur principal: Modérateur

Acteur secondaire : Modèle de Détection d'Émotions

Précondition : Pendant une réunion

Le modérateur visualise les émotions détectées à l'aide du Modèle de Détection d'Émotions.

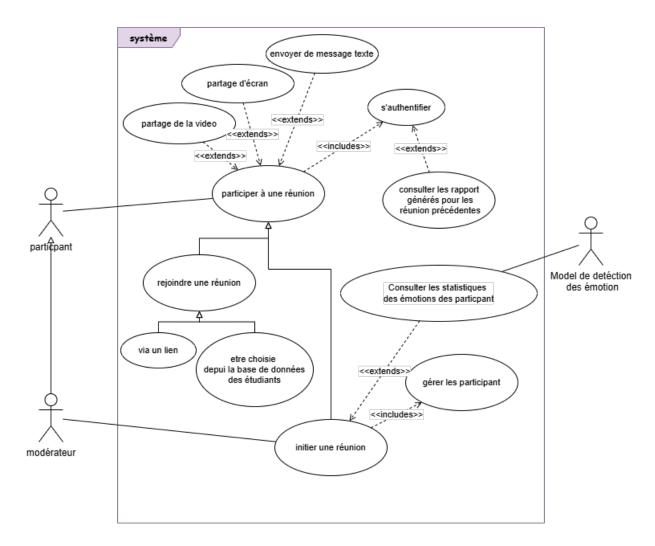


FIGURE 2.1 – Diagramme de Cas d'Utilisation

Diagrammes de Séquence

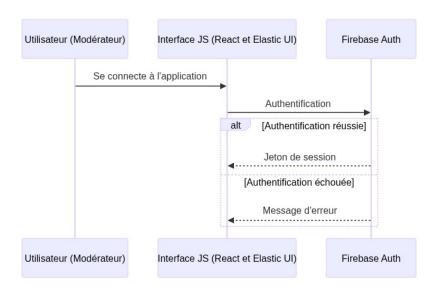


FIGURE 2.2 – Diagramme de Séquence (connexion et authentification)

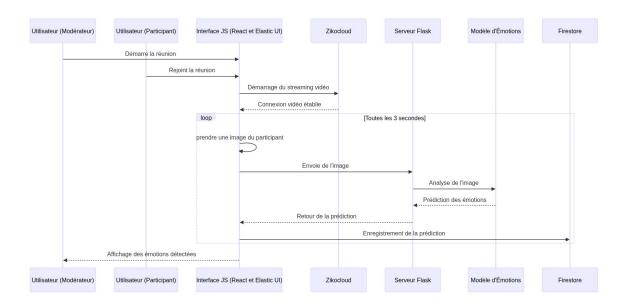


FIGURE 2.3 – Diagramme de Séquence (déroulement de la réunion)

Diagrammes de Déploiement

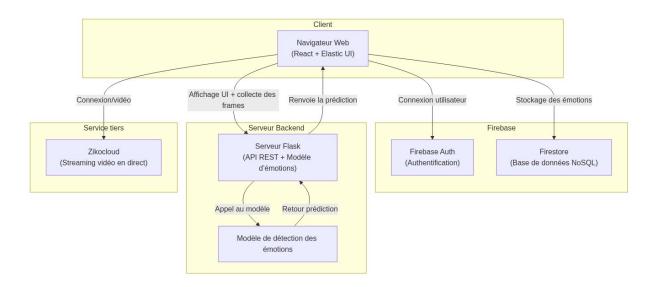


FIGURE 2.4 – Diagrammes de Déploiement

CHAPITRE 3

IMPLÉMENTATION

3.1 Introduction

Ce chapitre présente les détails techniques de l'implémentation de notre système de visioconférence avec détection des émotions. Il couvre les choix technologiques, la structure générale de l'application, les interactions entre les différents composants, ainsi que les principales fonctionnalités réalisées.

3.2 Choix Technologiques







ZEGOCLOUD



Figure 3.1 – Technologies utilisées dans le projet

Le système est conçu comme une application web moderne combinant plusieurs technologies côté client et serveur :

- React : pour la création de l'interface utilisateur dynamique.
- Elastic UI: pour la conception d'interfaces accessibles et modulaires.
- Firebase Authentication (FireAuth) : pour la gestion de l'authentification sécurisée.
- **Firestore** : base de données NoSQL pour stocker les prédictions d'émotions et les historiques.
- **Zegocloud**: pour l'intégration de la visioconférence (caméra, micro, écran).

— Flask (Python): pour le backend chargé de la détection d'émotions.

3.3 Architecture Générale du Système

Le système repose sur une architecture client-serveur intégrant également des services tiers. L'utilisateur interagit principalement via l'interface web, tandis que les traitements d'analyse émotionnelle sont assurés côté serveur.

3.3.1 Composants du Système

— Frontend (React + Elastic UI):

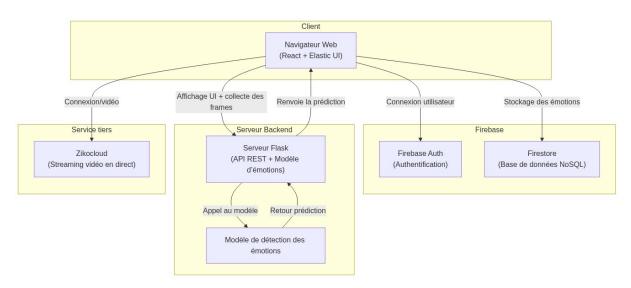
- Authentification utilisateur via FireAuth.
- Intégration avec Zegocloud pour les réunions.
- Capture des images toutes les 3 secondes et envoi vers le serveur Flask.
- Affichage des prédictions d'émotions en temps réel.

— Backend (Flask + modèle ML) :

- Réception des images au format base64.
- Traitement avec un modèle d'apprentissage automatique entraîné pour la reconnaissance des émotions.
- Renvoi des résultats (ex : joie, tristesse, colère, etc.) au frontend.

— Firestore :

- Enregistrement des résultats par utilisateur et par réunion.
- Accès possible par le modérateur pour consultation ou analyse.



3.4 Fonctionnalités Implémentées

— Authentification des utilisateurs.

- Lancement et participation à une réunion.
- Partage de caméra, d'écran, et messagerie instantanée via Zegocloud.
- Envoi périodique d'images vers le backend pour analyse.
- Affichage en temps réel des émotions détectées.
- Accès du modérateur aux statistiques des émotions des participants.

3.5 Conclusion

Cette implémentation permet d'assurer une expérience de visioconférence enrichie par l'analyse des émotions en temps réel, avec une architecture modulaire, légère et efficace. Dans le chapitre suivant, nous présenterons les résultats obtenus et les tests réalisés.

CHAPITRE 4

RÉSULTATS

4.1 Connexion

4.1.1 Authentification

Grâce à l'intégration de Firebase, l'utilisateur peut s'authentifier facilement via son compte Google.

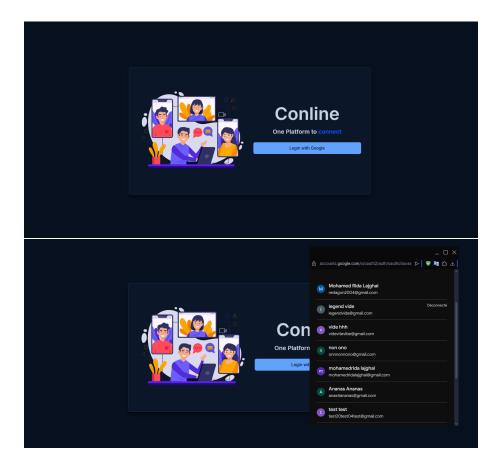


Figure 4.1 – Interface d'authentification Google via Firebase

4.1.2 Page d'accueil

L'interface d'accueil est simple, épurée et intuitive, facilitant l'accès aux fonctionnalités principales.

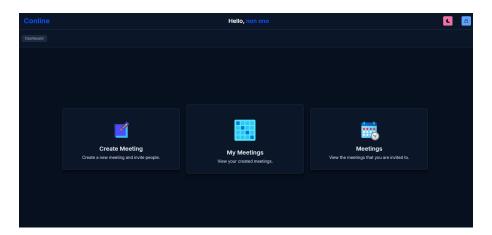


Figure 4.2 – Page d'accueil de l'application

4.2 Gestion des Réunions

4.2.1 Création d'une Réunion

L'utilisateur peut lancer une nouvelle réunion à partir de la page dédiée.

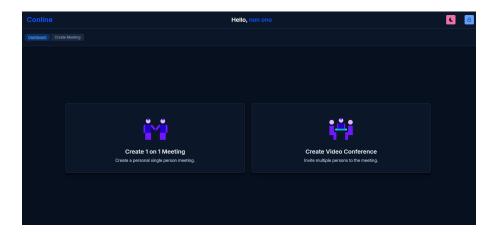


FIGURE 4.3 – Page de création de réunion

Réunion individuelle (1 à 1)

L'utilisateur peut créer une réunion individuelle en :

- saisissant un nom de réunion,
- sélectionnant un participant,
- programmant une date et une heure.

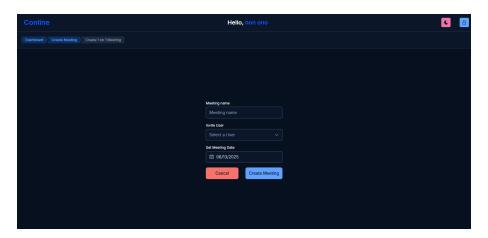


Figure 4.4 – Création d'une réunion individuelle

Vidéo-conférence

L'utilisateur peut également créer une réunion en mode visioconférence avec plusieurs participants.

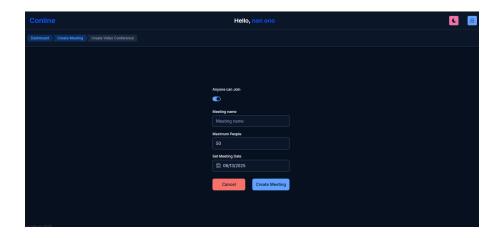


FIGURE 4.5 – Création d'une visioconférence

4.2.2 Rejoindre une Réunion

Les participants peuvent rejoindre une réunion via un lien d'invitation ou à partir de leur tableau de bord.

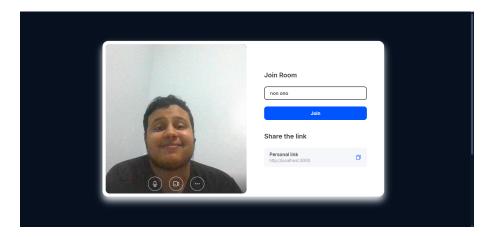


Figure 4.6 – Interface pour rejoindre une réunion

4.2.3 Déroulement d'une Réunion

En tant que participant

- Le participant peut :
- activer ou désactiver sa caméra et son microphone,
- partager son écran,
- envoyer des messages via le chat intégré.

En tant qu'organisateur

L'organisateur dispose des mêmes fonctionnalités qu'un participant, mais peut également :

— consulter les statistiques émotionnelles en temps réel fournies par le modèle d'IA intégré.

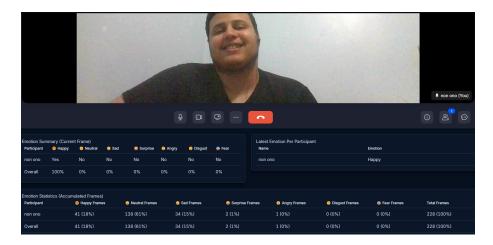


FIGURE 4.7 – Déroulement d'une réunion en cours

4.3 Historique des Réunions

4.3.1 Réunions créées

L'utilisateur peut consulter toutes les réunions qu'il a créées, les modifier si nécessaire, et accéder à leurs statistiques.

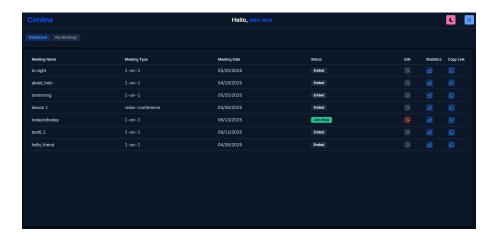


FIGURE 4.8 – Liste des réunions créées

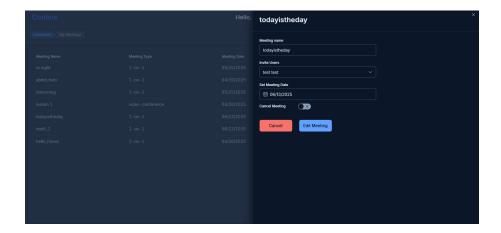


FIGURE 4.9 – Modification d'une réunion

Consultation des statistiques

L'organisateur peut visualiser les statistiques des émotions collectées lors des réunions passées.



Figure 4.10 – Consultation des statistiques d'une réunion

4.3.2 Réunions rejointes

Cette section permet à l'utilisateur de consulter l'historique des réunions auxquelles il a été invité.

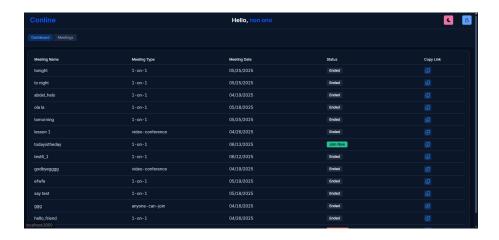


FIGURE 4.11 – Réunions rejointes

CHAPITRE 5

CONCLUSION ET PERSPECTIVES

Bilan du Projet

Ce projet avait pour objectif de concevoir et développer une application web interactive permettant d'organiser des réunions en ligne, tout en intégrant une fonctionnalité innovante de détection des émotions en temps réel à l'aide d'un modèle d'intelligence artificielle.

Réalisations techniques

- Mise en place d'une interface utilisateur moderne et réactive en **React.js** avec la bibliothèque **Elastic UI**.
- Authentification sécurisée via **Firebase Authentication**.
- Intégration de **Zikocloud** pour la diffusion vidéo en direct.
- Envoi périodique (toutes les 3 secondes) d'images extraites du flux vidéo vers un serveur **Flask** contenant un modèle de détection des émotions entraîné en **Python**.
- Stockage des prédictions dans **Firestore** pour une visualisation post-réunion.
- Interface permettant à un **modérateur** de créer une réunion, de gérer les participants, et de consulter les statistiques émotionnelles en direct et a posteriori.

Points positifs

- Intégration fluide entre plusieurs technologies hétérogènes.
- Traitement en temps réel des émotions avec un retour rapide vers l'utilisateur.
- Interface utilisateur ergonomique et facile à utiliser.
- Bonne modularité du code et facilité d'évolution future.

Difficultés rencontrées

- Problèmes de synchronisation entre le flux vidéo et l'envoi des frames.
- Limitations imposées par certains navigateurs lors de l'accès à la caméra.
- Optimisation du modèle d'émotion pour répondre en temps réel sans surcharge serveur.
- Intégration des technologies de streaming avec les composants front-end.

Améliorations possibles

- Ajout d'un système de notifications intelligentes basé sur les émotions dominantes.
- Implémentation de dashboards analytiques plus avancés pour les statistiques d'émotions.
- Passage à un modèle de détection plus performant (ex. : avec vision transformers).
- Déploiement de l'application complète sur un environnement cloud sécurisé (Firebase Hosting, Google Cloud Functions...).

Conclusion

Ce projet a permis de mettre en pratique des compétences en développement web, en intégration de services cloud, et en machine learning. Le résultat est une application fonctionnelle, intuitive et dotée d'une valeur ajoutée innovante grâce à l'analyse des émotions. Malgré certaines contraintes techniques, les objectifs principaux ont été atteints, offrant une base solide pour des évolutions futures.

Travaux Futurs

Ce projet constitue une base solide pour des améliorations et extensions futures. Plusieurs pistes peuvent être envisagées pour enrichir l'application et renforcer son utilité pédagogique.

- Détection multimodale : ajouter l'analyse des expressions textuelles (chat) et du ton de la voix (audio) pour une compréhension plus fine de l'état émotionnel des participants.
- Modèle plus adapté : entraîner un modèle de détection émotionnelle sur des données annotées spécifiquement pour le contexte éducatif, en ciblant des états comme la confusion, l'ennui, ou le niveau de concentration.
- Dashboards intelligents : concevoir des tableaux de bord dynamiques pour permettre aux enseignants de visualiser l'évolution émotionnelle d'un groupe ou d'un étudiant donné, sur une session ou dans le temps.
- Suggestions et alertes : proposer un système d'alertes ou de recommandations pour aider l'enseignant à adapter son discours ou son rythme lorsque des signes de désengagement sont détectés.

Ces perspectives visent à transformer l'application en un outil d'analyse pédagogique intelligent, au service des enseignants et des apprenants.