Chapitre 01 :

Concepts des méthodes d'authentification du Saint Coran

1. Introduction

Avec l'énorme quantité d'informations téléchargées par les utilisateurs tous dans le monde entier.
L'internet est la bibliothèque centrale que tous les êtres humains dans le monde utilisent pour échanger des informations.
Cependant, L'un des problèmes majeurs du Web actuel est qu'il existe n'y a pas de contrôle sur le type, la nature ou les détails des informations téléchargées sur le Web par n'importe quel utilisateur.
L'authenticité et l'intégrité des documents en ligne sont de plus en plus importantes obtenir crucial que de nombreuses organisations mettent leurs documents en ligne.
Fournir des installations qui permettent aux propriétaires et aux téléspectateurs de documents de pouvoir s'assurer que leurs documents sont authentiques et non altérés est extrêmement important.
Dans ce chapitre, nous allons définir le Saint Coran et montrer les principales méthodes d'authentification.
Nous allons aussi parler du traitement en langage naturel du texte coranique et enfin mentionner quelques domaines d'application du Saint Coran.

2. Le Saint Coran :

Le Coran comprend 114 chapitres de différentes longueurs (appelés sourates).
Chaque chapitre se compose de plusieurs versets (appelés Ayat).
Le nombre de verset la longueur de chaque verset diffèrent d'un chapitre à l'autre.
Le nombre total de versets dans le Coran dépasse 6000 versets.

2.1 Définition du Saint Coran :

Le Coran est la parole d'Allah est un livre de conseils pour tous les peuples du monde entier.
Le Coran couvre tous les aspects de l'économie à la politique, la science, à l'astronomie.
Il contient de nombreuses informations du début à la fin des temps.
Le Coran se compose de 30 Juz, qui a été descendu à Le prophète Muhammad SALLA ALLAHOU ALYHI WA SALLAM dans 23 ans.
Le Coran est un enregistrement du temps et une source importante pour retracer le développement de l'islam de La Mecque à Médine.
Jusqu'à présent, le Coran est considéré comme le livre le plus fiable au monde.
Par conséquent, la communauté musulmane considère le Coran comme la plus haute littérature en langue arabe.

2.2 Les styles d'écritures du texte Coranique :

En raison des différentes caractéristiques des lettres arabes comme les signes diacritiques (symboles de ponctuation), kashida (lettres étendues) et autres symboles, il est écrit et disponible dans différents styles comme Kufi, Naskh, Thuluth, Uthmani, etc.
Mais généralement nous avons utilisés Uthmani et texte brut (Plain).
Ces symboles Tajweed ne sont pas utilisés dans l'écriture arabe et sont uniques au Saint Coran. Ces signes diacritiques et Tajweed sont utilisés avec chaque lettre du Coran.
La figure 1.1 montre un exemple de verset du Coran contenant des signes diacritiques complets et du Tajweed.[1]

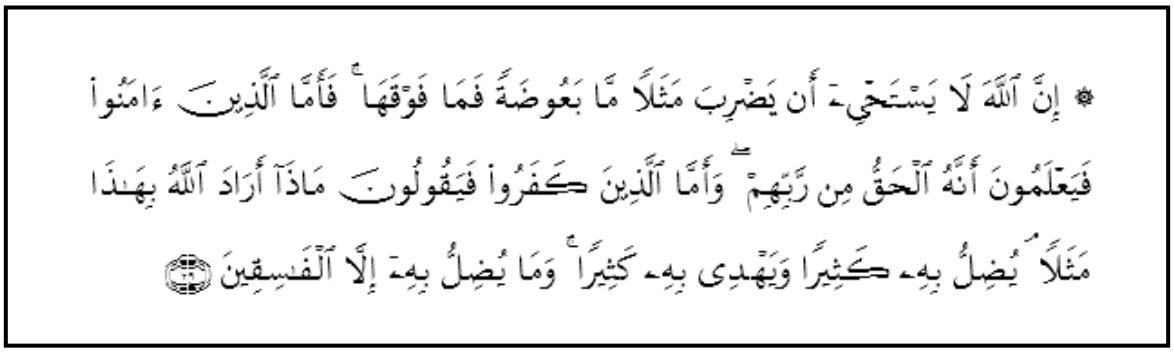


Figure 1.1 : Un exemple de verset du Coran contenant des signes [2]

2.3 Les récitations du Saint Coran :

La récitation du Saint Coran ou Qiraat est la manière de prononcer le Coran.
Les diverses récitations ne modifient pas le message du Coran ni ne prouvent qu'il existe des «versions différentes» du Coran.
Le tableau ci-dessous énumère les cinq principales citations du Saint Coran.

Récitations	Pourcentage	Usage
HAFS	**95%**	La majorité du monde musulman
WARSH	**3%**	Algérie, Maroc, Tunisie, Afrique de l'Ouest et Soudan.
HISHAM	**1%**	Yémen (partie)
QALUN	**0,7%**	Libye, Tunisie et Qatar (partie)
AL-DURI	**0,3%**	Soudan (partie) et Afrique de l'Ouest

Tableau 1.1 : Les cinq principales citations du Saint Coran.

3. Les méthodes principales d'authentification du saint Coran

Il existe différentes méthodes et approches pour résoudre problèmes spécifiques. Dans certains cas le type de format utilisé est crucial.
Par exemple le format d'image qui consiste en données brutes dépend de certaines techniques alors que le texte format qui se compose de données binaires dépend de son propre ensemble de techniques.
L'analyse de la recherche a révélé le manque de bonne direction car les chercheurs ont combiné le format de texte avec le format d'image et ainsi créée une confusion considérable. En outre, les études ne mentionnent pas si une approche particulière fonctionne pour différents formats.
Basé sur ces observations il est proposé de regrouper le corpus existant de recherche sur la préservation et la vérification du contenu coranique dans format basé sur l'image, format basé sur audio/vidéo et texte format. [2]

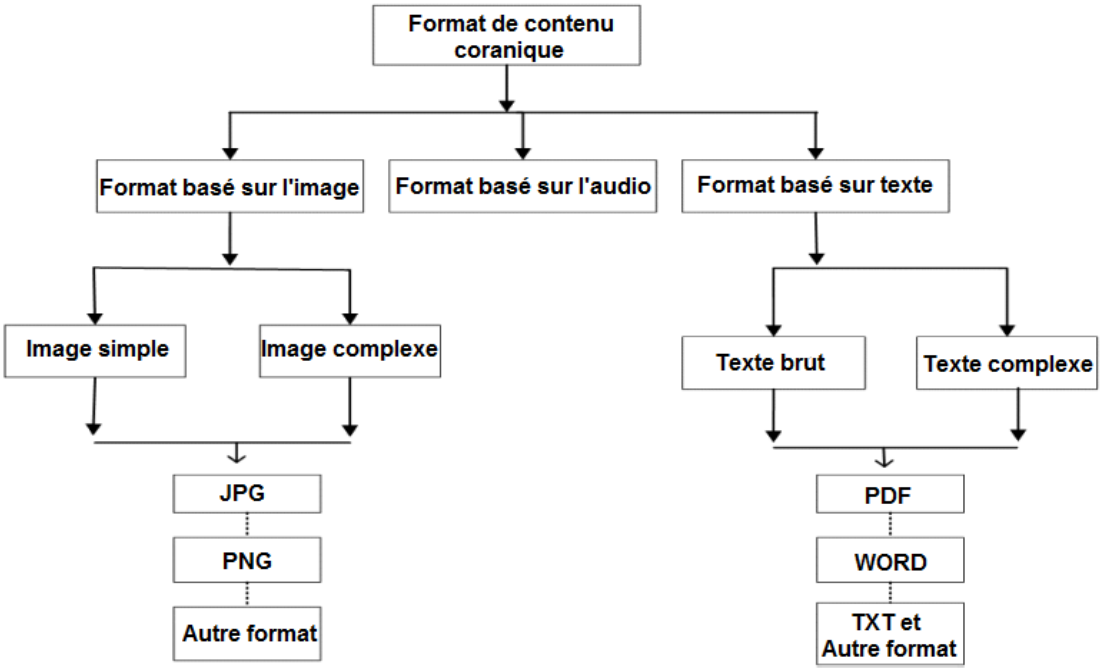


Figure 1.2 : Classification basée sur le format. [2]

3.1 Authentification du Saint Coran basée sur l'image

Il y a beaucoup d'images du Coran disponibles en ligne.
Afin de simplifier la classification, L'image contenue a divisée en deux sous-catégories, à savoir image brute et image complexe.
Une image simple constitue une image claire et nette sans beaucoup de détails de couleur alors qu'une image complexe est une image avec plus de détails de couleur et beaucoup de symboles
La Figure 1.3 : présente deux types d'images. Cette classification est faite sur la base du fait que différentes techniques sont utilisées dans le traitement d'images. Des images simples et complexes les images sont disponibles dans différents formats comme JPEG, TIF,GIF, etc.
Cette zone relève du domaine du traitement d'images, et il existe différents types de techniques utilisées pour vérifier l'intégrité des images. [2]

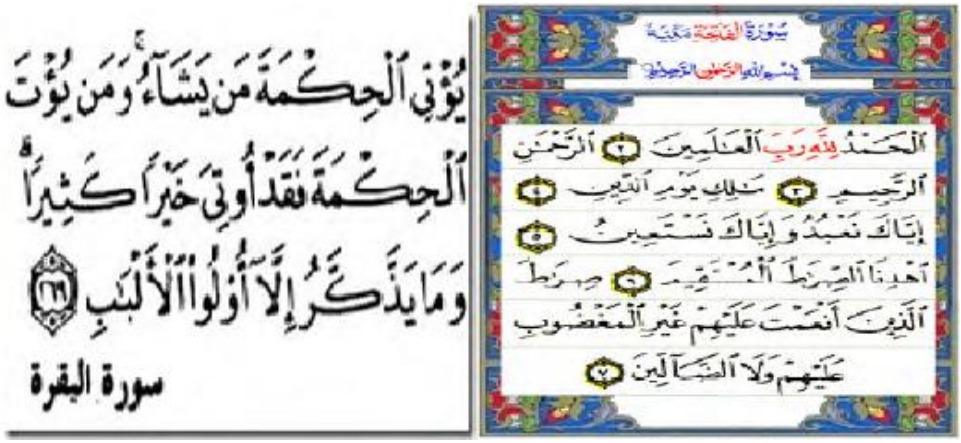


Figure 1.3 : La différence entre image simple et image complexe [2]

Les auteurs affirment que la protection du contenu et la protection du droit d'auteur constituent les deux principaux défis tout en traitant du contenu sensible en ligne.
Le Zero watermarking est identifié comme l'approche la plus prometteuse pour la vérification et l'authentification du contenu.
Un document a être authentifié est intégré à une séquence de données spécifiques obtenues à partir d'un logo en filigrane.
La limite de taille de mot pour laquelle la clé doit être générée n'est pas mentionnée.
Selon l'opinion, l'auteur calcule la clé pour l'ensemble du document, ce qui est similaire au processus de hachage.
Cette clé est utilisée par l'autorité de certification pour l'authentification lors de la phase de décodage.
La réclamation des auteurs en ce qui concerne la nouveauté de cette approche semble plutôt mal étayée.
Selon l'algorithme, les valeurs Unicode de tous les caractères sont calculées et ajoutées pour produire une somme.
Puis le bit de parité est ajouté à cette somme et la clé est générée. Par conséquent, la différence entre hachage et Zero watermarking pas claire. [3]

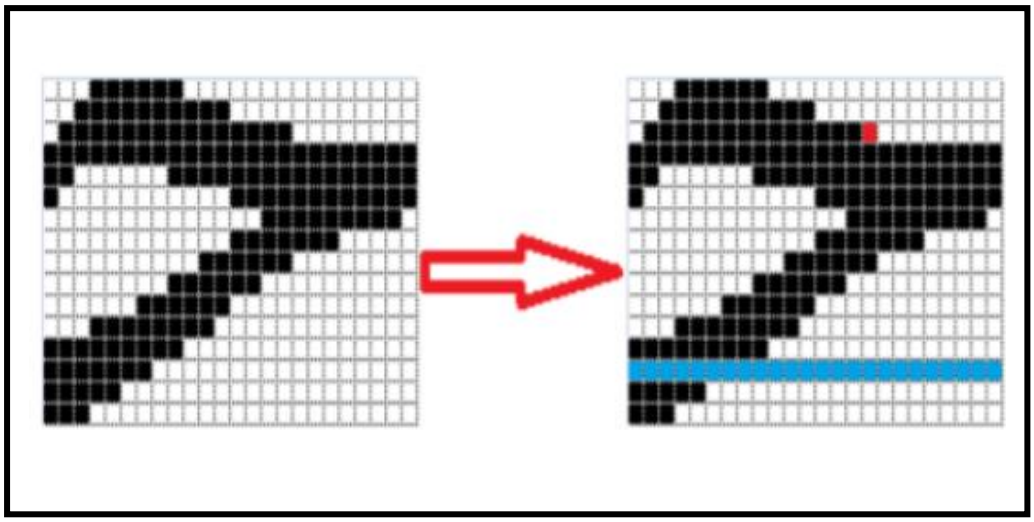


Figure 1.4: Le filigrane basé sur l'image, qui change le pixel de l'image en filigrane intégré

3.2 Authentification du Saint Coran basée sur un texte

La plupart des internautes préfèrent copier du texte coranique à partir d'une source et collez-le sur des sites de médias sociaux ou autres blogs en ligne.
Cette approche peut être appelée approche binaire et est l'une des approches les plus faciles et conviviales.
Donné la complexité des scripts arabes, le contenu du texte est divisé en sous-catégories, texte brut et texte complexe.
Par texte brut, on entend une simple écriture arabe sans aucun symbole.
Par texte complexe, on entend les écritures arabes avec l'ajout de signes diacritiques et de symboles, comme illustré à la Figure 1. 5.
Les différents formats disponibles sont le format Word, le format PDF, etc. [2]



Figure 1.5 : Différents styles d'écriture du Coran [8]

Par rapport au format basé sur l'image, moins de recherches ont été fait sur du texte.
Cela peut être dû à la complexité de la saisie de texte.

Comme le montre la figure 1.5, tous les styles présentés diffèrent par la façon dont les signes diacritiques et les autres propriétés écrites, comme les points, sont disposés.
La plupart des locuteurs natifs de l'Arabe n'ont pas besoin de signes diacritiques lisez le Coran, comme indiqué sur la figure 1.5 (D).
Cependant, il est essentiel que les locuteurs non natifs utilisent ces signes diacritiques pour le réciter et le comprendre correctement.
Par exemple, les diacritiques de base du Coran est illustrée à la figure 1.6. Si les diacritiques sont mal placés dans un verset, toute la signification de ce verset est modifiée.
Cependant, la plupart des approches existantes liées à l'authentification des textes du Coran Numérique (DHQ) éliminent ces signes diacritiques pour améliorer les résultats de récupération.
Une liste des symboles diacritiques et des symboles Tajweed (l'ensemble des règles relatives à la récitation) indiquent où arrêter la récitation et sont illustrées aux figures 1.6 et 1.7, respectivement. [8]

Futtha	◌
Thummah	◌
Tenween Futtha	◌
Tenween Thummah	◌
Tenween Kusrah	◌
Kusrah	◌

Tableau 1.2 : Diacritiques de la langue arabe.

continuer c'est mieux	صلے
doit arrêter	س
la garniture est meilleure	قلے
doit continuer	لا

Tableau 1.3 : Symboles de Tajweed.

Dans le style Uthmani, la lettre entourée d'un cercle rouge s'écrit différemment du style simple.
Il est également à noter que "alif" écrit en langage clair est simple et sous une forme standard. En général, dans le cas des vers d'Uthmani, un petit alif apparaît sur la lettre mim pour exprimer le son de la lettre alors que le script en clair ne l'inclut pas.
Cependant, les deux vers écrits en uthmani et en script clair sont corrects. Puisqu'aucun algorithme existant ne peut authentifier différents styles d'écriture coranique à partir d'une vérité commune (ensemble de données créé et vérifié manuellement), une nouvelle approche est nécessaire pour authentifier différents styles à l'aide d'une base de données commune.
Il est clair qu'il n'existe actuellement aucune approche efficace pour authentifiez différents styles de textes coraniques à l'aide d'une base de données commune.
Par conséquent, l'objectif de cet article est de proposer une approche pouvant résoudre les problèmes d'authentification des textes coraniques.
disponible en ligne qui sont écrits dans des styles différents. [8]

3.3 Authentification du Saint Coran basée sur un vocal

Cette partie porte sur l'authentification par filigrane audio et la protection de l'intégrité dans le contexte des données de parole transmises sur Internet dans un environnement d'apprentissage en temps réel.
La récitation du Coran arabe à distance est utilisée comme exemple d'étude de cas donnée sensible nécessitant une authentification robuste et des mesures d'intégrité.
L'approche de tatouage proposée ici est basée sur la détection des motifs de clé dans le signal audio en tant qu'entrée de l'algorithme l'application développée peut facilement être utilisée des deux côtés de la communication pour assurer l'authenticité et l'intégrité du signal de parole transmis.
Elle s'avère efficace pour de nombreuses applications d'apprentissage à distance nécessitant un traitement peu complexe en temps réel. [4]
Les signaux audios sous leur forme numérique sont facilement reproductibles sans distorsion. Des techniques de protection efficaces sont devenues essentielles.
Ce besoin a été renforcé par les nouveaux supports de distribution tels qu'Internet et par les méthodes de compression telles que les normes MP-3 et MP-4.
Le filigrane a été proposé pour résoudre ce problème. Le filigrane est un signal intégré au signal audio d'origine.
Le signal en filigrane contient ensuite des informations (données cachées pouvant être utilisées à de nombreuses fins).
Par exemple, les données cachées peuvent être utilisées pour identifier le propriétaire (protection du droit d'auteur permettant d'identifier la source des copies illégales, de vérifier l'intégrité d'un document, de faire un suivi).
Un signal dans un réseau et inclure des informations supplémentaires telles que le titre, l'auteur, les numéros de série, etc. Les applications du filigrane audionumérique sont nombreuses et comprennent : la protection du droit d'auteur, la protection contre la copie, la protection contre la falsification, le contrôle d'accès, les empreintes digitales, la gestion des droits numériques, authentification du contenu, contrôle des annotations et de la confidentialité, expertise des médias, amélioration de la communication, protection du contenu, filtrage du contenu, audit amélioré, localisation du contenu, etc.
Les contraintes qui doivent satisfaire à un schéma de tatouage audio dépendent de l'application. [5 - 6]

4. Traitement du langage naturel

Notre travail porte en particulier sur l'authentification de texte coranique.
Alors le traitement du langage naturel est important pour obtenir de bons résultats et assurer la vérification et l'authentification du texte Coranique.

4.1 Définition :

Traitement du langage naturel est un algorithme qui manipule une chaîne afin de maximiser son utilisation lors de la récupération de données.
Le traitement du langage naturel a deux rôles principaux à jouer dans le stockage et la récupération de grandes masses d'informations: fournir une interface conviviale et facile à apprendre aux systèmes de récupération d'informations et structurer automatiquement les textes de manière à ce que leurs informations puissent être traitées et récupérées plus facilement.

4.2 Traitement du langage naturel pour Rechercher dans le texte Coranique :

Le langage naturel a été impliqué dans de nombreux systèmes et jeux.
L'une des fonctionnalités du traitement automatique du langage consiste à supprimer les mots inutiles, les caractères spéciaux et même les symboles.
Cela rend l'interaction beaucoup plus interactive puisque lorsque l'utilisateur saisit des symboles, l'application reconnaît toujours le mot-clé requis. En second lieu, la séparation des mots permet à l'application de recevoir de nombreuses entrées de la part de l'utilisateur, telles qu'une phrase, ce qui permet à l'application de discuter avec l'utilisateur.
Le traitement automatique du langage implique également l'intelligence artificielle récupérée des données en fonction d'un mot-clé spécifique.
Outil de recherche de coran en utilisant le traitement du langage naturel, utilisez la fonction de rognage du traitement du langage naturel, séparez les mots et faites correspondre les mots-clés pour afficher la réponse correcte en fonction des entrées de l'utilisateur. [7]

5. Les domaines d'applications du saint coran

Cette section passe en revue les applications liées à la découverte du Coran en langage naturel.
L'examen est important pour comprendre le fonctionnement du système et comment il peut être amélioré pour les besoins futurs.

5.1 Les domaines d'applications du Coran sur les applications mobiles :

Il y a plusieurs des domaines d'application du Saint Coran , on va classifier ces domaines sur les systèmes d'exploitations des applications mobiles.

5.1.1 Application iQuran par Isysway :

Cette application est inventée pour aider l'utilisateur à lire le Coran.
Il est très interactif et comporte un signet pour la dernière récitation du Coran.
Cette application Coran d'Isysway est disponible au téléchargement gratuitement sur Playstore.
Il fournit également un mécanisme de recherche pour rechercher les versets dans chaque sourate que les utilisateurs veulent lire.
La fonction de recherche utilise un code fixe pour afficher le résultat de la recherche. Elle est similaire à la méthode traditionnelle de lecture du Coran, mais avec un téléphone portable. La figure 1.8 montre l'interface utilisateur graphique de l'application iQuran. [7]

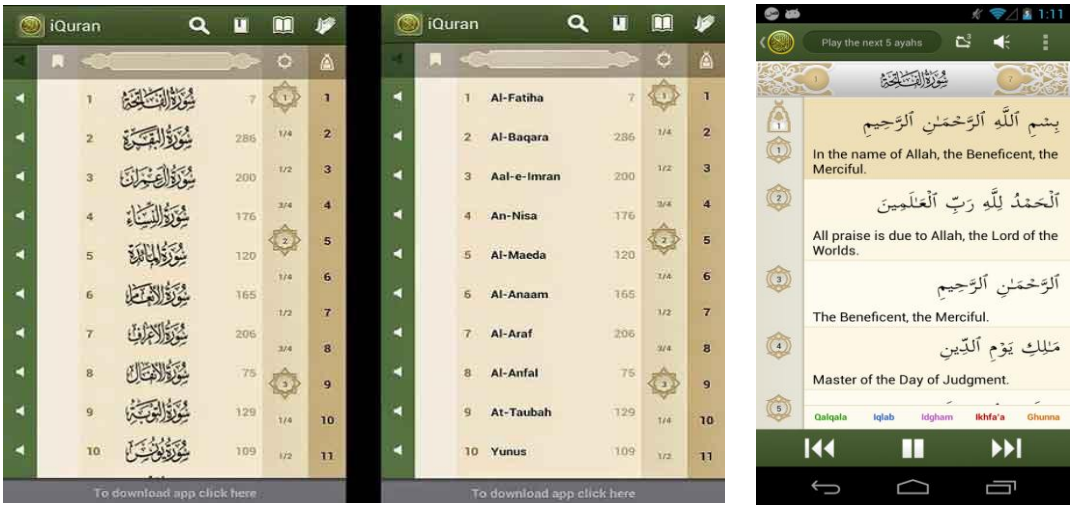


Figure 1.6 : Interface graphique d'iQuran.

5.1.2 Coran pour Android par Coran.com

Coran pour Android est une application source gratuite pouvant être téléchargé par les utilisateurs d'Android à l'aide de Google Play Store.
La fonctionnalité principale de l'application consiste à lire le Coran dans un smartphone.
L'application nécessite une connexion Internet pour que l'utilisateur puisse télécharger de l'audio pour l'application.
En raison de cette limitation, l'application a des problèmes pour arrêter l'audio du Coran lorsqu'il y a un appel sur le téléphone de l'utilisateur.
L'application dispose d'un mode nuit qui aide à réduire le stress des yeux lors de la lecture de textes à partir d'un écran lumineux dans une zone sombre.
Il utilise également des signets pour marquer le dernier verset de la lecture.
Cela permet de s'assurer que les utilisateurs lisent en se basant sur la continuité correcte du Coran.
La figure 1.9 montre l'interface du Coran pour Android application. [7]

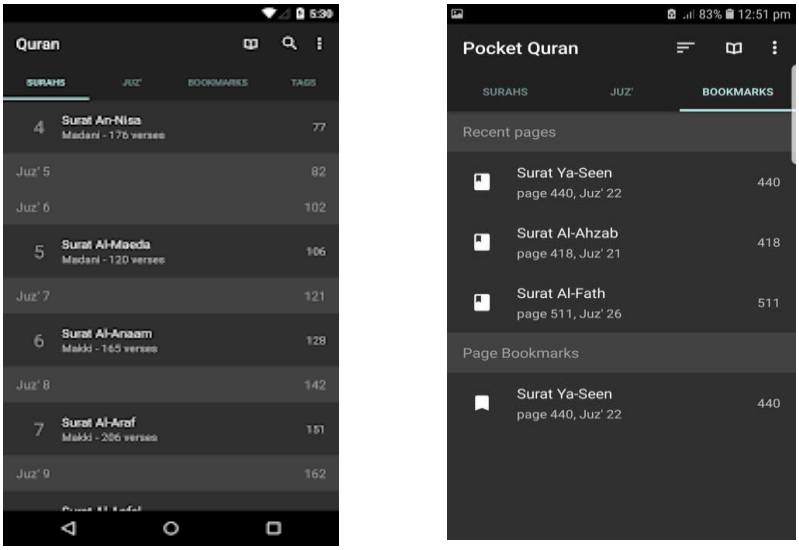


Figure 1.7 : Interface graphique du Coran pour Android.

5.1.3 Apprendre le Coran par Badr Interactive

Learn Coran est conçu pour les utilisateurs qui souhaitent apprendre à lire le Coran.
Étant donné que l'apprentissage du Coran est très difficile et que certaines personnes disposent de peu de temps, l'application est développée pour contrer ce problème.
Cette application aide l'utilisateur à apprendre le Coran de la prononciation à tajwid, l'audio aidant l'utilisateur à atteindre la prononciation correcte pour chaque verset. Bien que cette application soit gratuite, elle comporte une fonction premium que l'utilisateur peut obtenir lorsqu'elle est payée, telle qu'aucune publicité ni plus de didacticiel d'apprentissage sur le Coran.
La figure 1. 10 illustre l'interface utilisateur de l'application learn Quran. [7]



Figure 1. 8 : Interface graphique de PexPix.

5.2 Pour les sites web et ordinateurs :

Le Coran sur les sites Web est disponible, car nous pouvons voir qu'il existe de nombreux sites Web du Coran avec des styles différents et qu'il existe également des textes et des images et des fichiers audio et vidéo à lire pour jouer ou même télécharger.
On peut citer, par exemple https:/quran.com/ The noble Quran.
http://cran Muslim-web com/ The holy Quran.

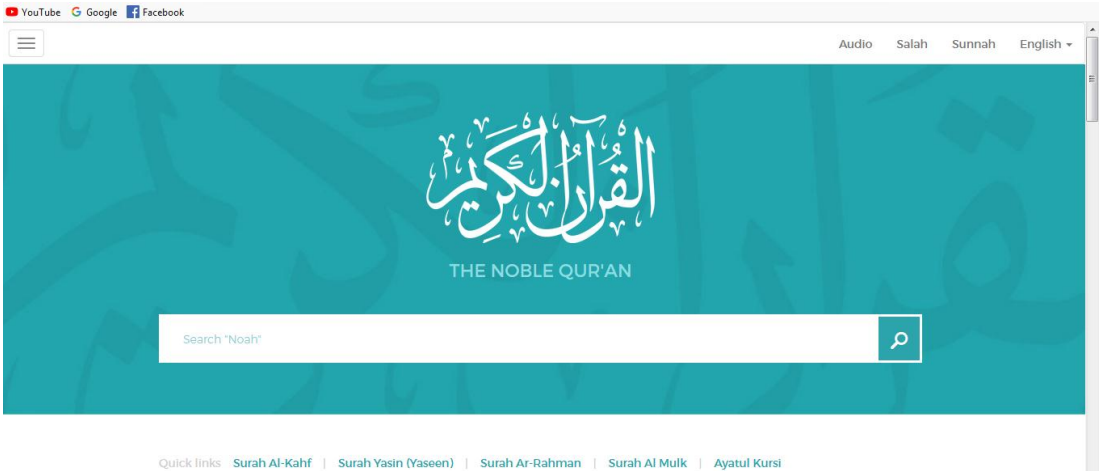


Figure 1. 9 : Interface graphique de web site The noble Quran.

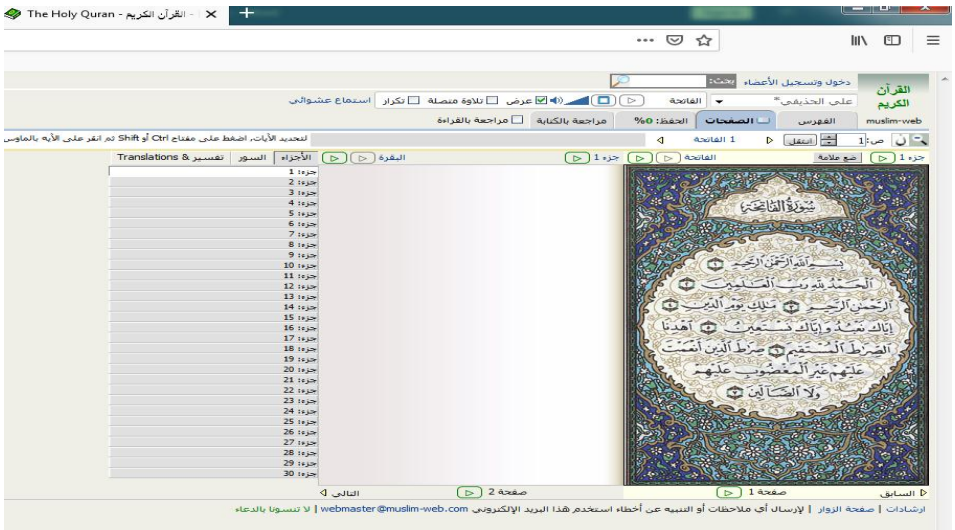


Figure 1. 10 : Interface graphique de web site The holy Quran.

Il existe de nombreuses applications du Coran exécutées sur un ordinateur avec les mêmes options que les sites Web ou les applications mobiles.

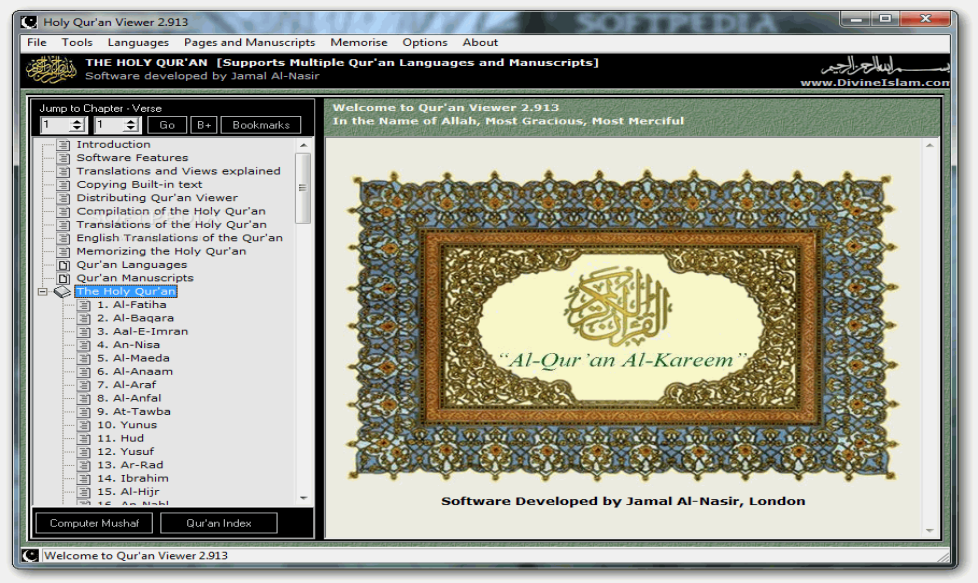


Figure 1. 11 : Un exemple d'une application du saint coran exécutée sur Windows

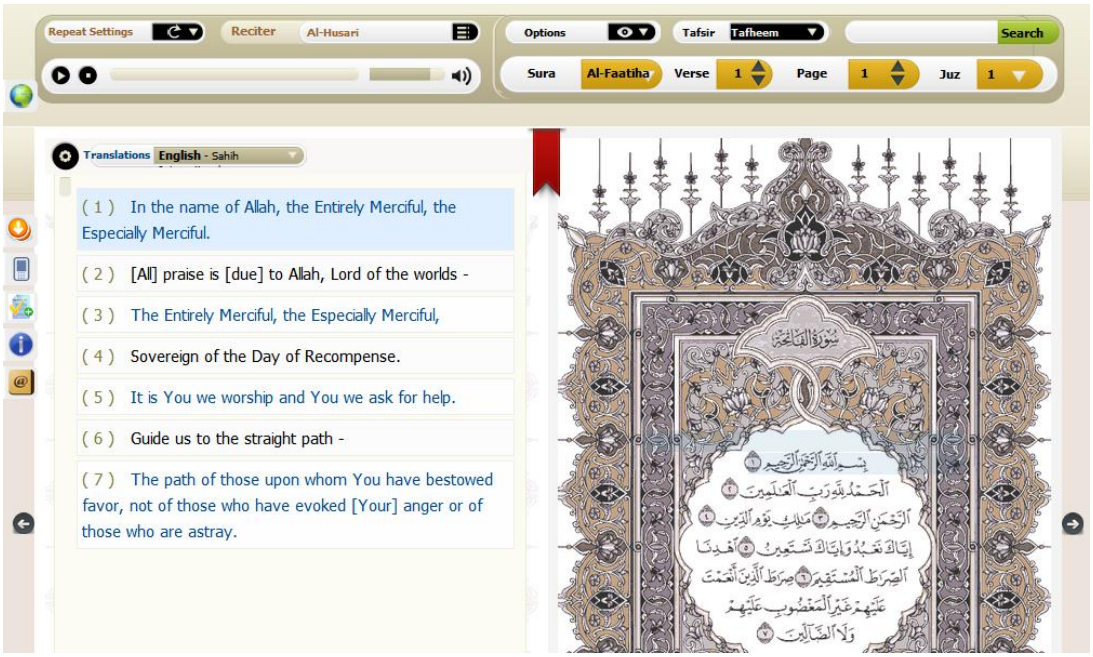


Figure 1. 12 : Un exemple d'une application du saint coran exécutée sur Mac

6. Conclusion

Nos travaux futurs porteront sur l'amélioration de la phase de vérification en travaillant sur les limites de l'approche proposée et l'étendre pour résoudre des styles plus complexes.
Il y a toujours beaucoup de problèmes à résoudre.
Tout d'abord, la précision de l'approche proposée doit être améliorée.
Deuxièmement, la disponibilité du Coran numérique dans différents styles est une autre recherche problème.
Il serait intéressant d'étendre l'approche proposée pour authentifier d'autres styles.
En outre, il sera intéressant de travailler à l'amélioration de la complexité temporelle de notre proposition approche et évaluer sa précision sur des grandes ensembles de données.