



**DRPI**

Direction Recherche  
Partenariats Innovation

## **AVIS DE PRESENTATION DE THESE EN SOUTENANCE POUR L'OBTENTION DU DIPLOME NATIONAL DE DOCTEUR**

**Monsieur Made Windu Antara KESIMAN**

Présentera ses travaux intitulés :

**« Analyse d'image de document des manuscrits balinais sur feuilles de palmier »**

Spécialité : informatique et applications

**Le 5 juillet 2018 à 16h30**

Lieu :

**Université de La Rochelle  
Pôle Communication, Multimédia et Réseaux  
Amphithéâtre  
44 Av. Albert Einstein  
17000 LA ROCHELLE**

Composition du jury :

<b>Mme BARNEY SMITH Elisa H.</b>	<b>Professeure, Boise State Université (USA)</b>
<b>M. BURIE Jean-Christophe</b>	<b>Professeur, Université de la Rochelle</b>
<b>M. DANTES Gede Rasben</b>	<b>Maître de conférences, Université Pendidikan Ganesha (Indonésie)</b>
<b>M. FISHER Andréas</b>	<b>Maître de conférences, Université de Fribourg (Suisse)</b>
<b>M. GRANGÉ Philippe</b>	<b>Maître de conférences, Université de La Rochelle</b>
<b>M.LLADOS Josep</b>	<b>Professeur, Université Autonome de Barcelone (Espagne)</b>
<b>M. OGIER Jean-Marc</b>	<b>Professeur, Université de la Rochelle</b>

### **Résumé :**

Les collections de manuscrits sur feuilles de palmier sont devenues une partie intégrante de la culture et de la vie des peuples de l'Asie du Sud-Est. Avec l'augmentation des projets de numérisation des documents patrimoniaux à travers le monde, les collections de manuscrits sur feuilles de palmier ont finalement attiré l'attention des chercheurs en analyse d'images de documents (AID). Les travaux de recherche menés dans le cadre de cette thèse ont porté sur les manuscrits d'Indonésie, et en particulier sur les manuscrits de Bali. Nos travaux visent à proposer des méthodes d'analyse pour les manuscrits sur feuilles de palmier. En effet, ces collections offrent de nouveaux défis car elles utilisent, d'une part, un support spécifique : les feuilles de palmier, et d'autre part, un langage et un script qui n'ont jamais été analysés auparavant. Prenant en compte, le contexte et les conditions de stockage des collections de manuscrits sur feuilles de palmier à Bali, nos travaux ont pour objectif d'apporter une valeur ajoutée aux manuscrits numérisés en développant des outils pour analyser, translittérer et indexer le contenu des manuscrits sur feuilles de palmier. Ces systèmes rendront ces manuscrits plus accessibles, lisibles et compréhensibles à un public plus large ainsi que pour les chercheurs et les étudiants du monde entier. Cette thèse a permis de développer un système d'AID pour les images de documents sur feuilles de palmier, comprenant plusieurs tâches de traitement d'images: numérisation du document, construction de la vérité terrain, binarisation, segmentation des lignes de texte et des glyphes, la reconnaissance des glyphes et des mots, translittération et l'indexation de document. Nous avons ainsi créé le premier corpus et jeu de données de manuscrits balinais sur feuilles de palmier. Ce corpus est actuellement disponible pour les chercheurs en AID. Nous avons également développé un système de reconnaissance des glyphes et un système de translittération automatique des manuscrits balinais. Cette thèse propose un schéma complet de reconnaissance de glyphes spatialement catégorisé pour la translittération des manuscrits balinais sur feuilles de palmier. Le schéma proposé comprend six tâches: la segmentation de lignes de texte et de glyphes, un processus de classification de glyphes, la détection de la position spatiale pour la catégorisation des glyphes, une reconnaissance globale et catégorisée des glyphes, la sélection des glyphes et la translittération basée sur des règles phonologiques. La translittération automatique de l'écriture balinaise nécessite de mettre en œuvre des mécanismes de représentation des connaissances et des règles phonologiques. Nous proposons un système de translittération sans segmentation basée sur la méthode LSTM. Celui-ci a été testé sur des données réelles et synthétiques. Il comprend un schéma d'apprentissage à deux niveaux pouvant s'appliquer au niveau du mot et au niveau de la ligne de texte.